The Dissertation Committee for Daniel Allen Clark Certifies that this is the approved version of the following dissertation:

Does Motivation Moderate the Effectiveness of Retrieval as a Learning Intervention?

Committee:

Marilla Svinicki, Supervisor

Daniel H. Robinson, Co-Supervisor

Erika Patall

Gary Borich

Anthony Petrosino

Does Motivation Moderate the Effectiveness of Retrieval as a Learning Intervention?

by

Daniel Allen Clark, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctorate of Philosophy

The University of Texas at Austin

May 2013

Dedication

To my parents, for your unconditional love and sacrifice.

Acknowledgements

I can do all things through Christ. I would like to thank my parents for imparting this belief through teaching and their amazing example. I am also thankful for my beautiful fiancé, Glynis, for loving me through this process. Special thanks also go out to the committee members and other faculty at The University of Texas at Austin. In addition, I would like to thank some of my fellow UT graduate students: Jason Crandall, Nate McVaugh, and Alyssa Reinhart.

Does Motivation Moderate the Effectiveness of Retrieval as a Learning Intervention?

Daniel Allen Clark, Ph.D.

The University of Texas at Austin, 2013

Co-Supervisors: Marilla Svinicki, Daniel H. Robinson

The effects of using retrieval as a study method have been found to occur across many contexts, such as in classrooms, with different age groups, and for non-verbal materials (Rohrer & Pashler, 2010). Even though researchers have suggested that this intervention be implemented on a widespread basis, studies to date have not investigated how the important variable of motivation could have an effect on retrieval as a learning intervention. This experiment investigated whether motivational variables would moderate the effect that retrieval has on learning. In this study, retrieval, extrinsic incentives, and intrinsic motivation positively affected performance. Causality orientations did not have an impact on performance or moderate the effect of the incentives. However, none of the included motivational variables moderated the effect of retrieval on learning. These results suggest that retrieval as a learning intervention is equally effective across different motivational conditions.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The cognitive strategy of retrieving encoded information from memory has received extensive empirical support as an effective learning intervention across many different types of learning, groups of individuals, and learning contexts (Roediger & Karpicke, 2006a; Rohrer & Pashler, 2010). The extent of this research has even led many researchers to suggest that retrieval practice should be widely implemented in educational settings (McDaniel, Agarwal, Huelser, McDermott, & Roediger, 2011; Pashler et al., 2007; Phelps, 2012). If this implementation is to be widespread, it may be important for researchers and practitioners to know how this intervention interacts with different individual and situational characteristics (Brewer & Unsworth, 2012). The current study investigated this question concerning motivation, both at the individual level (personal tendencies) and the group level (offering incentives) with the goal of improving the utility of the retrieval intervention in education settings.

Retrieval-Enhanced Learning

Retrieval as a method for learning has been studied intermittently in educational and cognitive psychology for many years. These studies have investigated the various affordances of retrieval (Butler, 2010; Carpenter, Pashler, & Vul, 2006), how retrieval works across different situational constraints (Bangert-Drowns, Kulik, & Kulik, 1991; McDaniel et al., 2011; Tse, Balota, & Roediger, 2010), and how retrieval may interact with individual characteristics (Bouwmeester & Verkoeijen, 2011; Brewer & Unsworth, 2012).

Many of the more recent studies concerning retrieval and learning are surprisingly similar to the first one that was published over 100 years ago. Abbott (1909) had her participants memorize nonsense word lists either by simply re-reading the words for 16 min. or by re-reading the words and attempting to retrieve them from memory for 8 min. Even though the participants were given the same

amount of time to study the information, the participants remembered the lists that included time for retrieval much more readily than the lists that did not include time for retrieval.

In the early years of retrieval practice research, many studies showed that retrieval is very effective for learning basic information such as word lists (Gates, 1922; Glover, 1989; Spitzer, 1939), yet newer conceptions of learning often assert that there is much more to the knowledge gained in educational settings than sets of facts that one must be able to recall (Mayer, 1996; Schallert & Martin, 2003). Fortunately, researchers who study retrieval have not solely focused on rote retrieval of facts. Retrieval has been found to improve the transferring of knowledge to different settings (Butler, 2010; Carpenter et al., 2006; Rohrer, Taylor, & Sholar, 2010), encoding of subsequent material (Pastotter, Schicker, Niedernhuber, & Buml, 2011; Szpunar, McDermott, & Roediger, 2008; Wissman, Rawson, & Pyc, 2011), as well as reducing forgetting (Carpenter, Pashler, Wixted, & Vul, 2008; Kornell, Bjork, & Garcia, 2011; Pansky, 2012). Although these laboratory studies have suggested the many affordances of using retrieval during learning, such findings are not always readily applicable to real-world settings (Mayer, 1996).

Although many of the discoveries concerning retrieval have occurred in the laboratory, researchers have provided evidence that retrieval improves learning in many different contexts (Roediger & Karpicke, 2006a). The effects of retrieval have been found across online studies, laboratory-based and classroom-based research (Bangert-Drowns et al., 1991; McDaniel, Roediger, & McDermott, 2007; Vojdanoska, Cranney, & Newell, 2010) and in a wide range of disciplines, such as mathematics, science, and government (Bangert-Drowns et al., 1991; Roediger & Karpicke, 2006b; Wissman et al., 2011). Learning across many different types of content, such as visuo-spatial materials (Carpenter & Pashler, 2007; Wheeler & Roediger, 1992), prose (Gates, 1922; Roediger & Karpicke, 2006b; Spitzer, 1939; Wissman et al., 2011), lecture (Butler & Roediger, 2007), and paired associates (Abott, 1909; Carrier & Pashler, 1992; Izawa, 1967, 1970) appears to be improved by retrieval. Several

different retrieval formats (e.g. multiple-choice, open-book, with feedback) have also been used to produce retrieval effects (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008; Fazio, Agarwal, Marsh, & Roediger, 2010; Rohrer et al., 2010). The effect of retrieval on learning also seems to occur across many different age groups (Gates, 1922; McDaniel et al., 2007; Roediger & Butler, 2011).

Providing feedback by showing the correct answer to the participant after a retrieval attempt is not absolutely necessary for retrieval to improve learning (Roediger & Karpicke, 2006a), but it does enhance the retrieval effect (Butler & Roediger, 2008; Vojdanoska et al., 2010). Retrieval has been found to be the most effective when the act of retrieval is difficult; for example, short answer questions are usually more effective than multiple-choice questions in enhancing learning (Kang, McDermott, & Roediger, 2007). Even when retrieval involves accessing content that is readily available (like an open-book test), it improves learning (Agarwal et al., 2008).

Importantly, retrieval has been found to work in classroom settings in addition to laboratory settings (Bangert-Drowns et al., 1991; Butler & Roediger, 2007; McDaniel et al., 2007; McDaniel et al., 2011; Roediger, Agarwal, McDaniel, & McDermott, 2011; Spitzer, 1939; Vojdanoska et al., 2010). In a meta-analysis, Bangert-Drowns et al. (1991) found that retrieval improved student performance on exams as much as one half of a standard deviation. In addition to improving overall achievement, retrieval-enhanced learning may also improve the calibration between students' expected and actual grades (Koriat & Bjork, 2006; McCabe, 2011).

Not surprisingly, retrieval does not always improve learning relative to re-studying. One important caveat is that the timing of the retrieval intervention and the criterion task has to be arranged carefully (Roediger, 2008). If the criterion task occurs shortly after the initial encoding of the material, retrieval is not more effective than re-studying (Roediger & Karpicke, 2006b). Similarly, if the retrieval task occurs shortly after the initial encoding of the material so that the content may be retrieved easily from working memory, retrieval does not appear to improve long-term learning

relative to re-studying (Karpicke & Roediger, 2007). Other evidence also suggests that retrieval may have the effect of improving the recall of only the items that the individual is able to recall during retrieval, whereas re-studying may improve the recall of all of the content (Kornell et al., 2011). Despite all of this evidence of the efficacy of this intervention from the cognitive research, barriers continue to exist concerning more widespread implementation of this learning intervention.

Barriers to Implementation

Despite the extensive evidence of how retrieval improves learning, barriers continue to keep this method from widespread implementation. One of these barriers is that instructors might assume that students already use retrieval while studying for exams. Unfortunately, Karpicke et al. (2009) found that students generally do not use retrieval while studying. Not using retrieval during study may negatively impact performance. Therefore, Einstein, Mullet, & Harrison (2012) suggested that instructors should take on the role of (1) educating their students concerning the effect of retrieval on learning and (2) incorporating regular quizzing as part of their curriculum. For this intervention to be effectively implemented on a widespread basis by professional educators, it may be important to know how the intervention may impact different individual and situational characteristics.

Individuals who underperform on criterion tasks may benefit from retrieval more than higher performers (Bouwmeester & Verkoeijen, 2011). Similarly, individuals who score higher on tests of episodic memory and problem-solving may benefit less from retrieval interventions (Brewer & Unsworth, 2012). Tse and Pu (2012) suggested that individuals with lower working memory and test anxiety scores benefited from retrieval more than those with low working memory and high test anxiety. Tse and Pu also reported that test anxiety was not related to the effect of retrieval among individuals with higher working memory. Although these results are potentially useful, they are based

on only three studies, so there is much that is not known about how individual factors relate to the effect of retrieval.

Motivation and Retrieval-Enhanced Learning

One important individual difference is motivation. Student motivation has been found to be a moderating variable that may have an impact on academic achievement (Becker, McElvany, & Kortenbruck, 2010; Dickhauser, Reinhard, Diener, & Bertrams, 2009; Logan, Medford, & Hughes, 2011; Singh, Granville, & Dika, 2002). Currently, only two known studies have examined the interaction of a motivational variable on retrieval-enhanced learning. These two studies presented some contrasting and puzzling findings, which research in the field of motivation may help explain some of the results.

Clark, Crandall, and Robinson (2012) conducted an experiment in which some students were given an incentive to perform well on a test that took place a week after they watched a video lecture concerning computer security. These students were told that receiving a particular score on the final test (and only the final test) would allow them to leave the laboratory after about ten minutes of testing, whereas participants who did not earn the specified score would be required to stay longer and take follow-up tests. In contrast to many earlier studies, the group that retrieved the material via a test did not outperform the group that simply viewed a transcript of the lecture. In addition, the incentives did not have the effect of improving the performance of the groups that received them. There was an interaction effect of condition (retrieval vs. transcript) and incentive. For the participants who had engaged in retrieval, those who did not receive an incentive outperformed the group who had received an incentive. Thus, the incentive appeared to serve as an overjustification for the retrieval effect.

In contrast Kang and Pashler (2012) conducted three separate experiments investigating the effect of motivation on learning through retrieval practice. In their first two experiments, the authors used a monetary incentive in order to motivate students to remember the Swahili-English word

translations in conditions of retrieval or re-studying. Kang and Pashler used a between-groups design and a different incentive in their final experiment. Similar to Clark et al. (2012), the participants were told that better performance would reduce the amount of time they spent in the laboratory. The groups that received the incentive outperformed the groups that did not on the final recall test. Kang and Pashler also reported that there was no interaction between the incentive and the study method (retrieval or re-studying). In light of both of these studies, it seems that there are unanswered questions concerning the effect of incentives on retrieval. Some of these questions may be answered through better understanding of the impact of motivation on these effects.

Motivation, Self-Determination, and Rewards

Examining the effect of rewards on motivated behavior is the primary objective of Edward Deci's (1971) Cognitive Evaluation Theory (CET). This theory is an empirical cognitive theory concerned with finding conditions in which motivation is affected by different types of incentives (Ryan & Deci, 2000). Research in CET has shown in many ways how external rewards can have the effect of undermining motivation (Deci, Koestner, & Ryan, 1999; Ryan, 1982). Rewards seem to have this undermining effect when the incentive has the effect if threatening one's perception of competence (Vallerand & Reid, 1984) or autonomy (Ryan, 1982).

CET was the work of Deci and Ryan, whose work later led to the larger macro-theory of motivation called Self-Determination Theory (SDT). Ryan and Deci (2000) suggest that human functioning works best when humans are living in concert with their natural tendencies. However, these natural tendencies are more than drives for food and water; they involve an inherent tendency towards growth and self-expression (Ryan, 1995). SDT also includes four mini-theories in addition to CET. One mini-theory, Causality Orientations Theory (COT), suggests that individuals have personality traits that may predispose them to seek differentially rewarding environments. Some prefer autonomous environments that are fairly unrestrained,

whereas others may prefer controlled environments with clear goals and rewards. Still others may have an impersonal orientation, feeling mostly blown about by the breezes of life (Deci & Ryan, 1985). Although all five of the SDT mini-theories are related to the current study, COT has particular relevance because it includes variables which are at the level of individual personality and not accounted for in the earlier mentioned experiments concerning motivation and retrieval-enhanced learning.

Causality Orientations and Individual Characteristics

Since COT has been conceptualized as a theory of personality tendencies, studies have shown that individual causality variables seem to be related to other psychological personality variables. Individual measurements of *autonomy* causality orientations (ACO) are positively correlated with self-esteem (Deci & Ryan, 1985), extraversion, openness (Olesen, Thomsen, Schnieber, & Tnnesvang, 2010), ease of concentration, perceived control, and competence (Wong, 2000). On the other hand, measurements of *control* causality orientations (CCO) are related to external locus of control, being publically and privately self-conscious, exhibiting hostility, type-A personality behaviors (Deci & Ryan, 1985), and neuroticism (Olesen, 2011). CCO measures also are negatively related to agreeableness (Olesen, 2011), and academic achievement (Wong, 2000). Deci and Ryan found that measures of the *impersonal* causality orientation (ICO) were positively correlated with measures of social anxiety, self-derogation, depression, public and private self-consciousness, hostility, fear, shame, guilt, and an external locus of control. Olesen et al. also found that ICO measures were positively correlated with neuroticism and negatively correlated with extraversion. Despite all of these relationships with personality characteristics, Olesen et al. (2010) also reported a factor analysis that suggested the constructs from COT were empirically distinct from the other personality characteristics measured. Although these measures appear to be valid as personality characteristics, these constructs may be more valuable if they demonstrate an effect on behavior.

Causality Orientations and Motivation Contexts

Deci and Ryan (1985) suggested that causality orientations would be associated with environmental events, so that similar behaviors and cognitions could be found amongst individuals with similar causality orientations. Koestner and Zuckerman (1994) found that individuals with high scores on ACO measures did not allow negative feedback to impact their subsequent intrinsic motivation, whereas those with high CCO measures experienced a bolstered intrinsic motivation in the face of negative feedback. Finally, Koestner and Zuckerman also suggested that individuals with high ICO scores suffered reduced intrinsic motivation in the face of negative feedback. In another experimental study, Hagger and Chatzisarantis (2011) found that ACO participants did not suffer a reduction in intrinsic motivation when they were given a reward. CCO participants did suffer a reduction in intrinsic motivation when they were given a reward. Beyond simple characteristics, it appears that COT variables tend to moderate how individuals react to situations as well.

From the perspective of SDT, the participants in the Clark et al. (2012) and one of the Kang and Pashler (2012) studies could have been experiencing a reduction in intrinsic motivation. The reduction could have been the result of a combination of individual personality factors (causality orientations) and situational factors (task-specific intrinsic motivation and characteristics of the reward). Unfortunately, no evidence of this possibility was collected for these previous studies. In addition, it is possible that some of these factors may aid in explaining why these two studies had broadly discrepant findings, with one finding a negative effect of incentives on performance and the other finding a facilitative effect. Therefore, the current study was designed to examine the larger effect of motivation on retrieval-enhanced learning, as well as to explore possible empirical explanations for the observed effects from these previous studies.

Since retrieval as a study method is such an effective intervention for improving memory, reducing forgetting, and improving the transfer of knowledge, many researchers have suggested

widespread implementation in educational settings (Roediger & Karpicke, 2006a; Rohrer & Pashler, 2010). Although this intervention has been found to be effective across many different contexts (Bangert-Drowns, Kulik, & Kulik, 1991; McDaniel et al., 2011; Tse, Balota, & Roediger, 2010), only two studies have investigated how individuals with differing levels of motivation may be assisted or impaired by retrieval-enhanced learning (Clark et al., 2012; Kang & Pashler, 2012). Unfortunately, these two previous studies only included extrinsic motivation (incentives) and had somewhat contradictory results.

Therefore, many questions still remain to be answered: How do extrinsic incentives moderate the effect of retrieval on learning? Does intrinsic motivation moderate the effect of retrieval on learning? Finally, do individual causality orientations moderate the effect of retrieval on learning? The current study has been designed to shed light on these questions. Knowing the answer to these questions could help educators and instructional designers to understand if using retrieval as a learning intervention would be particularly helpful depending on the motivation of the learners.

The next chapter of this dissertation will include a more thorough review of the existing literature concerning the affordances of the retrieval intervention. Past research concerning motivation will also be discussed from the perspective of Self-Determination Theory. These reviews of past literature will then culminate in discussing the current study.

Chapter 2. Motivation and Retrieval-Enhanced Learning

In light of the foregoing introduction, the proposed study will investigate how motivation may moderate the effect of retrieval on learning. Many different studies have experimentally investigated retrieval-enhanced learning (Rohrer et al., 2010). However, current educational research in motivation is largely of the self-reported correlational variety (Hsieh et al., 2005). The overall goal of the current investigation is to improve the empirical knowledge and practical utility of both motivation and retrieval constructs by studying how motivation might impact learning from retrieval.

The Effect of Retrieval on Learning

Most scientists trace the systematic study of memory back to Ebbinghaus (1885/1913). However, knowledge about methods of memory may be even older. The first mention of the concept of memory or recall in literature may be in the book of Deuteronomy in the Old Testament of the Bible, probably written around 1406, B.C. After Moses had presented a set of commands to the people of Israel, he said:

> These commands that I give to you are to be upon your hearts. Impress them on your children. Talk about them when you sit at home, and when you walk along the road, when you lie down and when you get up. Tie them as symbols on your hands and bind them on your foreheads (Deuteronomy 6: 6-9)

Many years later in the third century before the common era, Aristotle wrote in *Parva Naturalia* that "[E]xercise in repeatedly recalling a thing strengthens a memory" (Hammond, 1902; Roediger & Butler, 2011) . Both of these historical accounts suggest

that humans may at some level recognize intuitively the value of retrieval in learning and memory.

In the modern industrial era, intuitive knowledge has been discounted in favor of empirical research. Much of the early empirical work concerning memory was concerned with how the brain remembered information (Abott, 1909). The first empirical work concerning the positive effect of recall on learning was published by Abbott concerning her work in the psychology labs at the University of Illinois (Roediger & Butler, 2011). Abbott wrote that in watching participants memorize in the lab, she had noticed a tendency for them to "[M]omentarily turn away from the material that is before him [sic] and repeat it to himself without external aid (p. 159)." As a result, Abbott designed an experiment to study the "economy" of learning using both studying and recall. Her participants were told to memorize one list of words or random syllables per day in the lab. For one set of word lists, the participants were given 16 minutes to memorize the words. Another set of word lists was given in which the participants were directed to learn the words in 8 minutes and spend the other 8 minutes recalling them. For the third set of word lists, participants only saw the words for a total of 4 minutes, although this presentation was interspersed with recall attempts that took the remaining 12 minutes. Four hours later, the participants were asked to recall the lists of words. Although Dr. Abbott did not have the advantage of modern statistical methods, she suggested, "recall adds as much or more than the extra eight minutes of imprinting." Knowledge of this phenomenon would be further supported by subsequent studies (Tulving, 1967). In the final conclusion, Abbott said, "[T]he factor of recall is always an aid in the learning

process"(p. 177). This type of experiment became the prototype for many other similar studies concerning the effects of retrieval on learning.

Some years later, Gates (1922) performed a similar experiment, in which he used retrieval or re-studying methods to teach school children in Oakland, California. In these experiments, the children in grades 3-8 were told to study a nonsense word or a brief biographical paragraph he was holding on a card and to "never look away from the paper to see if you can say the words…never say a single word unless you are looking at it." (p. 30) The children were also taught a recall method Gates called "recitation (p. 1)." This method involved the children looking at a material, then looking away and trying to recite the word silently. Gates found that this retrieval method was much more effective in the amount of material that was learned than the earlier reading-only method. This effect occurred for both the nonsense words and the biographical paragraphs. Gates' study also reported that the effect of his retrieval study method appeared to actually increase with the delay of the test on which the children were measured. In the initial test, Gates reported that the retrieval studying method produced scores that were double the amount of the reading method, but that difference increased to quadruple when the delay was longer (three or four hours). Gates concluded that the process of retrieval is an effective study method that could be readily applied in educational contexts.

Spitzer (1939) performed a study with 6th graders in Iowa, using materials that were more aligned with normal educational practice, as well as different recall schedules. Many of the previous studies had used lists of nonsense words or short paragraphs for the content to be learned, a procedure that would continue in this type of research. However,

Spitzer's study involved the students' reading 600-word passages. In addition, Spitzer's retrieval method involved multiple-choice questions, instead of the recall methods that were used in the previous studies. Spitzer's study was unique because different recall schedules were utilized. One group was given recall tests on the same day, whereas other groups were given these tests the following day. The remaining groups were given the recall tests one, two, or three weeks after the initial reading. The results suggested that the recall study method was more effective when recall occurred within a week of the initial presentation of the material.

The results of these and other studies concerning the effect of retrieval on learning seem to have been enough for the research community to believe that the retrieval effect was a legitimate empirical intervention for improving learning (Pashler et al., 2007). Questions still remained concerning the extent to which retrieval may affect learning in the classroom.

The Affordances of Retrieval Interventions

Even though these classic studies have shown that the power of retrieval is very interesting, it is common for educators to suggest that there is much more to the knowledge gained in educational settings than sets of facts that one must be able to recall (Mayer, 1996). As a result, researchers have not solely focused on testing recall. In addition to improving recall, retrieval appears to; improve the transferring of knowledge to different settings (Butler, 2010; Carpenter et al., 2006; Rohrer et al., 2010) and encoding of subsequent material (Pastotter et al., 2011; Szpunar et al., 2008; Wissman et

al., 2011), as well as reducing forgetting (Carpenter et al., 2008; Kornell et al., 2011; Pansky, 2012).

Retrieval improves transfer of knowledge. Transferring knowledge is a diverse concept that involves applying gained knowledge in a novel situation (Salomon & Perkins, 1989). The initial studies concerning this phenomenon in relation to the retrieval effect included a fairly simple procedure for measuring transfer. Carpenter, Pashler, and Vul (2006) conducted studies in which participants learned pairs of words that were translations from German to English. After studying the word-pairs, the participants in the first study took a cued recall test or were given an opportunity to re-study the set of words. Cued recall tests involve being given one of the words and being asked to recall the other word in the pair. For example, if the word-pair was "Hund – Dog," then the cued recall test stem would be "Hund - ?" asking for the word "dog" to be entered by the participant. For this particular study, the outcome test was different from previous studies because it included cued recall tests, which in some instances resembled the first cued recall test, but in other instances offered the opposite word as a cue. For example, if the word-pair was "Hund – Dog," then the first test would say "Hund - ?," while the second test might say "Dog - ?" or "Hund - ?" Carpenter et al. reported that the participants who had taken a cued recall test outperformed the groups who had simply re-studied the word pairs on both types of cued recall outcome tests. Carpenter et al. concluded that retrieval improved the participants' memory for both the word they were seeking to recall and the word that was associated with it. In their second study, the same word pair task was used, but the first cued recall test given to the retrieval group did not include all of the word

pairs that were on the initial list. When the second test was given which did contain all of the initially studied word pairs, Carpenter et al. found that the retrieval group showed enhanced recall even for the words that were not even included in the retrieval task.

In a similar set of experiments, Butler (2010) used prose material from six different passages to assess the extent to which retrieval would benefit different types of knowledge transfer. By manipulating the outcome measures, Butler's experiments showed that retrieval improved the participants' ability to answer questions that were formatted differently than the initial test questions (Experiment 1b) and answer questions about knowledge domains that were not included in the initial test (Experiment 3).

In a different type of study, Rohrer, Taylor, and Sholar (2010) taught 4th and 5th graders to locate fictitious cities on a map in which all of the cities were connected by a road. After learning the locations, the retrieval group was required to try and recall the locations of each particular city before the actual location was revealed. The re-study group was simply required to write the location of the city as it appeared on the screen. One of the outcome tests included a standard test of whether or not the children remembered the locations of the cities on the map. There was also a transfer test that required a very different procedure. The transfer test involved the child looking at an unlabeled version of the map and answering questions concerning the navigation of the places on the map. For example, one question was "If you drove from Ross to Boyd, which cities would you drive through?" Rohrer et al. reported that the children who had earlier completed the recall test outperformed the other children on this novel task as

well. These studies suggest yet another benefit to retrieval: it seems to enhance transfer of knowledge into different settings.

Retrieval improves encoding of subsequent material. Another surprising affordance that comes with using tests as a learning activity involves the benefits that may arise in learning that occurs subsequent to the material that is being tested (Pastotter et al., 2011; Szpunar et al., 2008; Wissman et al., 2011). Szpunar et al. asked their college student participants to learn five lists of word pairs. Some groups were required to recall each list, while others re-studied the first four lists and were only tested on the last list. In a previous study using a similar procedure, the authors reported how this procedure had created an effect in which the participants in the re-study condition remembered the items from the first list well, yet the subsequent lists seem to have induced interference in their overall performance. This interference was called an "intrusion," because words from the list they knew well (list 1) were reported when they were being tested on subsequent lists (lists 2 – 5) (Szpunar, McDermott, & Roediger, 2007). In their subsequent 2008 studies, the authors found that retrieval may have had the effect of reducing the interference that had been occurring from learning several lists of words. Szpunar et al. (2008) called this phenomenon the "interim testing effect."

Subsequently, Wissman et al. (2011) found an interim testing effect using more complex prose materials instead of word lists. Using these materials, participants who had completed adjunct questions after sections of the text continued to outperform the participants who completed the recall at the end of the whole text. In one of their experiments, Wissman et al. included a group who did a non-retrieval activity (math

problems) at the same time that the interim testing group was completing the interim tests (Experiment 3). The results of this experiment showed that it was not simply taking a pause while reading the text that was accounting for the improvement in scores. Wissman et al. also noted the appearance of content intrusions taking place in the groups that did not take interim tests, yet suggested that intrusions may be more rare when content involves textual material instead of word lists which were used in that particular study.

In a rare and important experiment, Pastotter et al. (2011) essentially replicated the study design from the previous Szpunar et al. (2008) study with the addition of physiological measurement of brain oscillation from an electroencephalograph (EEG). Studies using EEG measurement had found previously that when studying lists of words, brain oscillations in the alpha frequency ranges measured by the EEG increased as the amount of words to be studied increased (Sederberg et al., 2006). Oscillations in the alpha range are characterized experientially by situations in which the working memory is highly engaged internally, such as when doing mental calculations (Palva & Palva, 2007). As would be expected from the aforementioned research, Pastotter et al. (2011) found that oscillations in the alpha range increased for participants in the re-study group and that individual increases in alpha range oscillations predicted poor performance on the final list. This suggests that as more word lists were re-studied, the working memory was working harder to accommodate the increased number of items to be remembered. By contrast, the EEG measurements for the participants in the retrieval group did not show the accompanying increase in alpha range oscillations across the five word lists. The authors suggest that the act of retrieval may lead the brain to effectively reset its own

memory encoding processes, which may enhance subsequent recall. This study may be significant in that it showed both a cognitive effect of the interim testing intervention as well as its physiological substrates.

Retrieval Reduces Forgetting. One of the most important aspects of retrieval-enhanced learning may be the effect that retrieval has on reducing forgetting (Carpenter et al., 2008; Kornell et al., 2011; Pansky, 2012; Runquist, 1983). For the purposes of clarity, there is a difference between improving recall and reducing forgetting. To illustrate, Rundquist (1983) conducted an experiment in which participants studied a list of 24 word pairs (Experiment 2). Following a two-minute long interpolating task, they used a cued–recall retrieval method on half of the studied items. Next, the participants did another set of un-related tasks, followed by a final test that included all 24 of the studied word pairs. For the purposes of this review, the final test would be the criterion test that measures the effect of retrieval on learning. As we have already seen in previous studies, retrieval does have a positive effect on this immediate recall test. In order to assess the effect of retrieval on forgetting, Rundquist used subsequent tests that occurred 2, 7, and 21 days after the initial presentation of the material. The results of Rundquist's study suggest that in these longer periods of time, the 12 items that were retrieved initially were still recalled at a higher rate than the 12 items that were not part of the retrieval study method initially. These results show that retrieval may increase recall and reduce the rate of forgetting.

More recently, Carpenter et al. (2008) conducted a similar experiment using word pairs that were translations between Swahili and English as well as obscure facts. Instead

of using a within-participants design like the previous study, Carpenter et al. tested the rate of forgetting across different groups for as much as 42 days. Also, Carpenter et al. tested the rate of forgetting using a curve-fitting method in the tradition of Ebbinghaus (1885/1913) and others (Rubin & Wenzel, 1996; Wixted & Ebbesen, 1997). Carpenter et al. reasoned that this type of analysis differs from an ANOVA approach in that it describes forgetting as a proportion loss of information learned instead of an absolute amount. Using the curve-fitting approach, the authors suggested that for the average participant's performance to drop below 25% would take an average of 3 days for the re-study group, whereas it would take 30 days for the retrieval group to fall below 25% retention.

From the simpler studies conducted over a hundred years in the past to the more complex studies using modern technology, retrieval has shown itself to be an empirically robust intervention with an extensive number of effects. However, interventions that work well in research labs can be very fragile and largely dependent on many different contextual variables (Roediger, 2008). These context differences may be very important when considering the potential of this intervention to be used in educational settings on a widespread scale.

Retrieval-Enhanced Learning Across Different Contexts

Throughout the years, researchers have provided evidence that the retrieval effect does improve learning in many different contexts (Roediger & Karpicke, 2006a).

However, new research is beginning to examine if retrieval has different effects for different types of individuals (Brewer & Unsworth, 2012; Tse & Pu, 2012).

Retrieval-Enhanced Learning in the Classroom

One problem that often occurs when it comes to applying cognitive research is that interventions are too fragile to create results in the real world (Roediger, 2008). Therefore, many researchers have investigated the ability of retrieval to translate from the laboratory to the classroom (Bangert-Drowns et al., 1991; Butler & Roediger, 2007; McDaniel et al., 2007; McDaniel et al., 2011; Roediger et al., 2011; Spitzer, 1939; Vojdanoska et al., 2010). However, some studies have been conducted in classrooms, but not actually as a part of regular classroom activity (Spitzer, 1939). In addition to improving achievement in the classroom, test-enhanced learning may also improve other aspects of the classroom experience (Koriat & Bjork, 2006; McCabe, 2011).

Bangert-Drowns et al. (1991) conducted a meta-analysis of 40 different classroom-only studies in which the teachers were testing the effect of frequent quizzing on students' performance. The studies were chosen because they were conducted in the classroom, used conventional classroom quizzes, were methodologically appropriate, and involved both groups receiving comparable instruction. The studies reported data from participants as young as 9th grade, but the majority of the studies were conducted in college classrooms. The disciplines from the studies were widely varied, including mathematics, statistics, veterinary medicine, psychology, accounting, physiology, law, geography, and more. Of the 40 studies selected, 35 of them investigated numerically the

question at hand. The authors reported positive retrieval effects in 29 of these studies and 13 were statistically significant. In the 11 studies in which the students in the non-retrieval groups did not take any tests at all, the groups who practiced retrieval performed an average of 0.54 standard deviations better on their course examinations. Four of the studies reviewed included the instructor surveying the students concerning their attitudes towards being tested before the exams. Among these four studies, the students who were in the testing groups on average rated their attitudes as more positive, with an average effect size of 0.59 standard deviations.

It is likely that the results of the four studies that found positive attitudinal effects of testing in the classroom may seem counter-intuitive to many readers. Surprisingly, regular testing may be one of a few different possible solutions for solving at least one pervasive problem in the classroom. This problem involves the tendency for students to be overconfident in their predictions of their performance (Grimes, 2002; Hacker, Bol, Horgan, & Rakow, 2000; Jensen & Moore, 2008; Nowell & Alston, 2007). In one of these studies, Grimes asked 253 students at Mississippi State University to predict their grades on an Economics exam two days before the exam occurred. After the exams were administered and scored, over 86% of the students overestimated the scores they would actually receive. When the students were asked to predict their scores after they had taken the test, 70% continued to overestimate the scores they would later receive.

Koriat and Bjork (2006) tested the possibility that testing might lead to students' making more calibrated predictions concerning their knowledge. While learning sets of word lists, students were required to study or attempt to recall what they had learned.

Before the final test on the words remembered, the students were shown the word pairs again and asked to predict the probability that the word pair would be successfully recalled. When the students had already attempted to recall the item, the correlation between their probability judgments and actual recall was much higher than when they had re-studied the items. Although this study was conducted in a laboratory setting, the point of the study was that testing has the potential to improve this common problem that exists in the classroom: student miscalibration of their own level of learning.

Since retrieval is an effective method to improve performance and reduce miscalibration, one might think students would naturally use retrieval when studying for their exams. Karpicke et al. (2009) studied the strategies that college students normally use when studying for exams. The students were instructed to list all of the methods that they use and which method they use the most often. Of the students surveyed, only 11% reported using recall as a method that they used to study, with only 1.1% (2 students) reporting that recalling was the method that they used the most often. There were other forms of testing that were more common, such as flash cards (40%) and practice problems (43%). Although most of the research reviewed has shown it to be a less effective method, 84% of the students reported using re-reading text materials as a study method. When given a choice between recall, re-studying, or another learning method, over 77% of the students chose options besides recall.

Similarly, McCabe (2011) found that a majority of surveyed college students predicted that re-studying would lead to better learning than retrieval (study 1). Thankfully, an additional study showed that students who were directly taught about the

empirical research concerning retrieval and learning were able to correctly deduce that retrieval would produce better learning than re-studying (McCabe, study 2). Unfortunately, McCabe noted that it remains unknown if learning about the benefits of retrieval makes an impact on the actual study behavior of the students.

In summary, using retrieval in the classroom context has been repeatedly shown to be an effective method for improving learning (Bangert-Drowns et al., 1991). Retrieval can also be used as a useful intervention that improves students' knowledge of their own learning (Koriat & Bjork, 2006), which might even improve an instructor's course evaluations (Isely & Singh, 2005). In order to take advantage of this improved learning calibration of retrieval-enhanced learning, new technologies may afford instructors methods to improve the usage of retrieval interventions in the classroom.

Retrieval Effects using Different Formats

In recent years, there has been a surge in educational use of new technology to transform assessments of learning (Beebe, Vonderwell, & Boboc, 2010). Two main affordances from this surge involve an increased ability to provide learning feedback (Adams & Strickland, 2012) and new types of learning assessments (Pellegrino & Quellmalz, 2011). Not surprisingly, researchers have investigated the interaction between retrieval interventions and learner feedback (Butler & Roediger, 2008; Kang et al., 2007) as well as different assessment formats utilized (Agarwal et al., 2008; Carpenter & Pashler, 2007).

When it comes to learning, the effect feedback is not always simple (Kulhavy & Stock, 1989). Feedback has been defined as "information provided by an agent regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007). In a synthesis of over 500 meta-analyses, Hattie and Timperley reported that feedback might have a larger impact on learning than prior cognitive ability, socioeconomic influences, and homework. With regards to retrieval interventions, question format and feedback are usually seen as moderators of the effects of retrieval.

Glover (1989) investigated the impact of question format on retrieval-enhanced learning by comparing the overall performance of participants who had performed different types of retrieval. Following the reading of prose passages, participants either re-studied or retrieved content in the form of free recall, cued-recall, or recognition tasks. The recall retrieval method involved simply recalling as much as possible of the essay. The cued-recall retrieval method consisted of sentences from the passages being presented with blanks that the participants were to fill in with words from the passage. The recognition retrieval method involved the participants reading a list of sentences in which half of them were exact copies of sentences they encountered in the passage. The participants were asked to choose which sentences they had read in the passage. Using these different types of retrieval methods, Glover reported that the free recall and cued recall retrieval methods resulted in better learning when compared to the control re-studying groups. However, the recognition format did not result in a retrieval effect.

Although multiple-choice tests are often considered tests of recognition, Glover's recognition test was much different. Roediger and Marsh (2005) applied the use of

multiple-choice tests as retrieval methods to assess whether taking a multiple-choice retrieval would result in a retrieval effect. There was some doubt, because it seems possible that incorrect answer options on the multiple-choice questions (called lures or distractors) would result in students choosing the wrong options on the final assessment. In two experiments, Roediger and Marsh found that multiple-choice retrieval methods could produce a retrieval effect. In addition, there was some truth to the theory concerning the effect of lures on final test performance. When the number of lures was increased, Roediger and Marsh found that the retrieval effect was indeed reduced.

More recently, Butler et al. (2008) examined a possible method in which the negative effects of the multiple-choice retrieval method could be mitigated. In this study, the participants read twelve different prose passages. For some of the passages, the participants did not retrieve the material. For others, they employed the multiple-choice retrieval method. For the remaining passages, they used the multiple-choice retrieval method and incorporated immediate or delayed feedback. The feedback was given in the form of the computer informing the participants if the given answer was correct, as well as the computer informing the participant of the correct answer to the question. The authors indicated that including feedback (despite whether it was immediate or delayed) appeared to enhance the positive effects of the multiple-choice retrieval method on later performance. Furthermore, including feedback also reduced the number of times that the participants chose one of the incorrect lures as the correct answer to the question (called "intrusions"). Based on these results, Butler et al. suggested that feedback is an important element for retrieval interventions.

Another test format that is used in education involves open-book tests. Agarwal et al. (2008) conducted a set of studies to investigate whether an open-book test can be used as a retrieval method to improve learning in the same way as the previously mentioned retrieval methods. In these studies, participants in the open-book conditions answered short-answer questions about prose material while they were able to look at the material. The other conditions involved re-studying the material or taking a short-answer test without being able to look at the text at the same time. Not surprisingly, the authors reported that participants who were able to refer to the text directly performed better on the initial test. However, the open-book testing conditions resulted in comparable learning to the closed-book testing conditions. In their second experiment, a similar effect occurred when participants in the re-study condition were allowed to re-study the text more than once.

Although these studies have shown that some differences occur between these different test formats (recall, multiple-choice, and open-book), they are actually fairly similar in that these test formats involve the measurement of verbal material. Carpenter et al. (2007) used non-verbal materials in order to investigate whether testing would continue to improve non-verbal learning. The undergraduates in the study were shown maps that contained drawings of land features (such as rivers and hills) and symbols (such as telephones and restroom symbols). In the re-study condition, the participants were given the opportunity to study the map and told that they would be tested. For the testing condition, the participants were shown a copy of the map with something missing from the earlier version. They were told to find the missing part on the first screen and

remember it. On the second screen, they were shown the right answer. In the end, the participants were given a blank sheet of paper on which to re-create what they could remember of the original maps. The authors of the study noted that the testing condition produced better learning than the re-studying condition even when using these non-verbal materials. These studies have shown that the testing effect appears to be a fairly robust phenomenon that seems to affect many different types of tasks in different contexts. However, one contextual variable that remains to be discussed involves how testing may interact with individual characteristics in order to impact learning.

Retrieval Effects and Individual Characteristics

With the diversity that is becoming increasingly common in educational contexts, a few researchers have become oriented towards investigating the effect that retrieval may have on individuals with different individual characteristics (Bouwmeester & Verkoeijen, 2011; Brewer & Unsworth, 2012; Tse & Pu, 2012). In the past, retrieval has been found to improve learning for different types of individuals, such as people in different age groups (Gates, 1922; McDaniel et al., 2011; Roediger & Butler, 2011).

For the majority of the studies on this particular topic, the convenience sample of college undergraduates has been used. As mentioned earlier, Gates (1922) found a testing effect using a sample of students in grades 1 – 8. The only age-related difference noted in the study involved the youngest group. Gates found that his recitation study method did not work well with the first grade students. He suggested that limiting the students' exposure to the material for these novice readers likely resulted in this negative effect for

this particular group. McDaniel et al. (2011) found retrieval enhanced learning for a sample of students in eighth-grade science classes. Bouwmeester and Verkoeijen (2011) studied a sample of students from third to sixth grades and did not report any age-based differences on the effect of retrieval.

Although it seems that retrieval-enhanced learning is largely not bound by age differences, Tse et al. (2010) found at least one difference in a study comparing the effect of retrieval on face-learning in middle-aged and older adults. In the first study, the participants were learning the names and occupations associated with pictures using testing or re-study methods. For the middle-aged adults, the expected testing effect occurred. Tse et al. indicated that the older adults benefitted more from the re-studying condition and showed more intrusions of incorrect answers in the retrieval condition. The authors attributed this difference to a possible decline in episodic memory amongst the older adults. In their second study Tse et al. performed a near replication of the first experiment, with the addition of providing feedback during the retrieval condition. The addition of feedback reversed the previous advantage of re-studying condition and reduced the amount of intrusion errors amongst the older adults. In the end, retrieval with feedback improved learning across all of the age groups in the second study.

In one of the first studies to investigate individual differences and retrieval-enhanced learning, Bouwmeester and Verkoeijen (2011) utilized a unique statistical analysis in order to investigate the possibility that individual differences might account for some of the differences observed in research concerning the testing effect. The result of this "Latent Class" analysis basically placed the children into one of three clusters. The

first cluster involved students who benefitted the most from the tested words. The second class involved students who did not benefit as much from testing. The third class involved students who seemed to have performed so well in the re-study condition that there was no difference between the testing and re-study words. The authors suggest that an individual's ability to effectively process a "gist trace" explains the ability to benefit from testing. Unfortunately, this characteristic seems to have been only measured by the same materials that were used to elicit the testing effect, so the reliability of this suggestion may be in question. Despite that weakness, the statement that individual characteristics may bear some responsibility for the testing effect may be fruitful for future research.

In order to improve upon this knowledge, Brewer and Unsworth (2012) tested undergraduate students' learning from testing as well as a host of other individual cognitive abilities. In addition to learning from re-studying or testing lists of Swahili-English translations, the participants took tests of working memory capacity, episodic memory, attentional control, and general intelligence. Similar to Bouwmeester and Verkoeijen (2011), Brewer and Unsworth reported that the individuals in the experiment appeared to differ with regards to the benefits that were afforded by the testing condition. In addition, Brewer and Unsworth investigated the ability of the aforementioned cognitive abilities to predict the benefit from testing as compared to re-studying. Of these abilities, the authors found that episodic memory and general intelligence were significant predictors of the testing effect. However, the direction of this prediction was negative. Specifically, individuals with lower scores on the general intelligence test

appeared to benefit more from testing than individuals with higher scores. Individuals with lower scores on the episodic memory measurements also appeared to benefit more from testing than individuals with higher scores on the episodic memory measures.

Tse and Pu (2012) conducted a similar study in which participants engaged in the same task as the Brewer and Unsworth (2012) study, with additional measures of working memory capacity and test anxiety included as predictors. Contrary to the results of the Brewer and Unsworth study, Tse and Pu found that working memory capacity was a significant predictor of the effect of retrieval. The participants with lower working memory scores benefitted more from retrieval than participants with higher working memory scores. In addition, the authors reported that the interaction between working memory and test anxiety was a significant predictor of the testing effect even when controlling for the main effect of working memory capacity. To explore this interaction, the authors performed separate analyses for different levels of working memory capacity. For the individuals with lower working memory capacity, test anxiety was a significant negative predictor of the benefit of retrieval. More specifically, individuals with lower working memory capacity and lower test anxiety benefitted more from retrieval than individuals with lower working memory capacity and higher test anxiety. For the individuals that had a higher level of working memory capacity, test anxiety was not related to the effect of retrieval. A similar interaction occurred when the authors were predicting the amount of intrusion errors made by the participants. Individuals with lower working memory capacity and lower test anxiety made less intrusion errors than individuals with lower working memory capacity and higher test anxiety. For the

individuals that had a higher level of working memory capacity, test anxiety was not related to the amount of intrusion errors.

Overall, these studies concerning individual factors and retrieval seem very preliminary when compared to other research in this field. Despite the fairly rigorous number of cognitive constructs measures included, the effects are not consistent and some of the differences found are relatively small. Another area of research concerning the effect of retrieval on learning that is quite sparse is theoretical reasons that may explain the phenomenon that has been observed in all of these years of research (Pyc & Rawson, 2009; Roediger & Butler, 2011).

Theories of Retrieval-Enhanced Learning

Three basic theories concerning the effect of retrieval on learning have been suggested across relevant research (Roediger & Butler, 2011). Spreading activation theory suggests that the act of retrieval provides the learner with an opportunity to create additional connections between different types of knowledge, therefore making the probability of retrieval higher because of the existence of more memory traces (Anderson, 1983). Transfer Appropriate Processing theory posits that retrieval has a positive effect on learning because the act of retrieval during the learning phase approximates the same retrieval process during the criterion learning task (Morris et al., 1977). Finally, Disuse theory suggests that humans have a limited retrieval capacity and that memories which are not retrieved after being learned are basically snuffed out by memories that are retrieved after being learned (Bjork & Bjork, 1992). In their recent

review article, Roediger and Butler suggested that these respective theories tend to have different aspects of the effect of retrieval that they explain well.

The basis of *spreading activation theory* is that knowledge forms into related cognitive units with associated traces that form like a spider web of connected strings (Anderson, 1983). This theory suggests that these traces have different levels of strength based on the number of times the unit is successfully recalled. With regard to the effects of retrieval, spreading activation theory suggests that retrieval allows for the activation of many different traces, therefore improving the probability of activating the correct response (Carpenter, 2009). Spreading activation theory seems to help explain the consistent finding that learning from retrieval is enhanced when retrieval is difficult (Kang et al., 2007; Kornell, Hays, & Bjork, 2009; Pyc & Rawson, 2009). However, research suggests that difficult retrieval is the most effective when the memory is successfully retrieved and retrieval may not be as helpful when it is not (Kornell et al., 2011).

Transfer appropriate processing theory suggests that learning may depend on the extent to which the method of learning approximates the method of representation of that which was learned (Morris et al., 1977). With regard to retrieval-enhanced learning, transfer appropriate processing theory views the act of retrieval during learning mostly as practice for the criterion task (Roediger & Butler, 2011). This theory has received support from studies that find that repeated retrieval is superior to retrieving information only once (Karpicke & Roediger, 2008).

In contrast to transfer appropriate processing theory, Kang et al. (2007) investigated the effects of different types of study retrieval formats on different types of criterion task formats. Some of the groups engaged in short answer retrieval while others retrieved information using a multiple-choice recognition format. When the criterion task was taken three days later, the participants also took either a short answer test or a multiple-choice test. Kang et al. reported that the participants' learning did not differ based on the method that was used on the final test, but different retrieval methods appeared to make a difference. Short answer retrieval was shown by Kang et al. to be the most effective retrieval method across different criterion test formats.

Disuse theory is somewhat similar to the aforementioned spreading activation theory, except disuse theory splits memory strength into storage strength and the retrieval strength (Bjork & Bjork, 1992; Roediger & Butler, 2011). Bjork and Bjork accounted for the effect of retrieval by suggesting that repeatedly viewing an item may only increase the storage strength of that memory, while retrieval of that item increased both the storage and retrieval strength of that particular memory. Also, disuse theory suggests that retrieval strength of a memory can be overcome by the competition of different memories for limited retrieval capacity (Bjork & Bjork). Roediger and Butler (2011) suggested that disuse theory can account for the spacing effects of retrieval. Spaced retrieval involves practicing retrieval periodically over longer periods of time, which has benefits in long-term memory (Carpenter & DeLosh, 2005).

Although these three theories may account for different aspects of the effects of retrieval on learning, several authors have noted that theories concerning this type of

research are currently inadequate (Pyc & Rawson, 2011; Roediger & Butler, 2011). One possible reason for this inadequacy could be that a variable is missing which has not been included in the research to date. This research has largely ignored the effect that motivation may have on retrieval-enhanced learning.

Motivation and Retrieval-Enhanced Learning

Student motivation has been found to be a moderating variable that may have a significant impact on academic achievement (Becker et al., 2010; Dickhauser et al., 2009; Logan et al., 2011; Singh et al., 2002) Currently, only two known studies have examined the interaction of a motivational variable on retrieval-enhanced learning. These two studies presented some contrasting and puzzling findings, which research in the field of motivation may help explain some of the results.

Clark, Crandall, and Robinson (2012) conducted an experimental study in which some students were given an incentive to perform well on a test that took place a week after they watched a video lecture concerning computer security. These students were told that receiving a particular score on the final test (and only the final test) would allow them to leave the laboratory after about ten minutes of testing, whereas participants who did not earn the specified score would be required to stay longer and take follow-up tests. The participants in the incentive conditions were informed of this prior to viewing the material during the first session and then they were reminded before they began taking the final test a week later. In addition, some of the participants engaged in retrieval of the

information with feedback while others were allowed to review a transcript of the video lecture. In contrast to many of the studies reviewed, the authors noted that the group that retrieved the material did not significantly outperform the group that viewed the transcript. In addition, the incentives did not have the effect of improving the performance of the groups that received them. However, the authors reported an interaction that occurred between the different conditions. For the participants who had engaged in retrieval, those who did not receive an incentive outperformed the group that had received an incentive. The authors noted that many different processes could explain this interaction. Given the results of the aforementioned Tse and Pu (2012) study, one possible explanation could be that the incentive induced the participants who received it to have more anxiety during the final test, which reduced their performance relative to those who did not receive an incentive. Unfortunately, there were no data collected that offered a plausible explanation, so Clark et al. suggested that further research might be able to untangle the interaction that was found in their study.

Kang and Pashler (2012) conducted three separate experiments investigating the effect of motivation on learning through retrieval practice. In their first two experiments, the authors used a monetary incentive in order to motivate students to remember the Swahili-English word translations in conditions of retrieval or re-studying. In their first experiment, one group of participants was offered either 5 cents or 30 cents for every word pair that they recalled correctly. Unfortunately, Kang and Pashler reported that offering 25 more cents did not have a significant impact on final performance. The second experiment also used a within-group design and compared how well participants

recalled Swahili word-pairs that had no reward or a reward of 25 cents. Similar to the first experiment, the authors reported that there was no difference between these incentives and final recall. In the third experiment, Kang and Pashler used a between-groups design and a different incentive. Similar to Clark et al. (2012), the participants were told that better performance would reduce the amount of time they spent in the laboratory. Using this incentive and a between-groups design, the groups who received the incentive did outperform the groups who did not on the final recall test. However, Kang and Pashler also reported that there was not an interaction between the incentive and the study method (retrieval or re-studying).

Even when not considering the contrast in results, these studies raise some very interesting questions. Why would an incentive to perform better not have an effect on performance? Why would an incentive that offered participants less time in the lab impact the final performance, whereas a monetary incentive did not? Finally, what other unmeasured factors may have had an impact on the results of both studies? In order to speculate regarding these questions, research on motivation must be considered.

Rewards and Behavior

Examining the effect of rewards on motivated behavior is the primary objective of Edward Deci's (1971) Cognitive Evaluation Theory (CET). This theory is an empirical cognitive theory concerned with finding conditions in which motivation is facilitated or undermined by different types of incentives (Ryan & Deci, 2000). CET posits that tangible external rewards usually have the effect of undermining motivation (Deci et al., 1999; Ryan, 1982).

Early research concerning motivation largely involved the manipulation of reinforcement from basic physiological needs such as food, water, air, or sleep (Greeno, Collins, & Resnick, 1996). This type of research began with animals, but it also postulated that human motivation could be traced back to the basic process of associating behaviors with physiological reinforcement (Thorndike, 1898). In the tradition of researchers of the time, Harlow (1950) used rhesus monkeys solving problems to learn about behavior and motivation. In contrast to normal experiments of the time, Harlow's monkeys were not solving the problems to receive a physiological reward like food or water. Harlow reported that the rhesus monkeys continued to attempt to solve the puzzles that they encountered on their own without the promise of a reward. In addition, the rate at which the monkeys learned the puzzles was almost identical to the rate of other monkeys who were solving the puzzles to obtain a reward. Subsequently, White (1959) reported similar evidence of this type of behavior occurring throughout many different species. This tendency to perform behaviors without any external contingency has been called intrinsic motivation (Deci & Ryan, 1980). This term can sometimes be best explained as the polar opposite of extrinsic motivation, which involves performing activities solely to receive some type of external reward.

In an experiment that was somewhat similar to what Harlow did with rhesus monkeys, Deci (1971) brought undergraduates to his lab to build configurations from a set of puzzle pieces in three sessions. During the first session, the participants were simply instructed to re-produce the designs printed on a piece of paper. During the second session, some of the participants were told that they would receive one dollar for

every configuration that was completed successfully within the time limit. When they arrived at the third and final session, those who had received money were told that the money had all been distributed so that no money was available for correct configurations. At one point during each session, the experimenter would leave the room and observe the participant through a one-way window. Intrinsic motivation measured in this experiment was the amount of time that the participant spent engaging with the puzzles while the experimenter was outside of the room. Deci reported that the participants who had received the monetary reward during the second session spent on average 50 seconds less of their free time on the puzzles than they had during the first session. Therefore, it seemed that offering the monetary reward might have had a negative effect on their intrinsic motivation.

Subsequently, Deci (1972) performed a similar experiment in which a monetary reward was compared with the effect of a reward that was strictly verbal. In these experiments, the participants were sometimes given praise for completing the configurations successfully. Deci's results tentatively supported his theory that verbal rewards would have a positive effect on the participants' intrinsic motivation. The results were tentative because it appeared that the verbal rewards only produced an effect for the males in the study. In a refreshing reminder that research actually occurs in the real world, Deci hypothesized that the verbal rewards were not as effective for the females in the experiment because the reward of interacting with an attractive male experimenter may have produced a positive effect in both of the conditions. Similar to the previous study, the monetary rewards continued to undermine the participants' intrinsic

motivation. Deci hypothesized two aspects of rewards that may affect intrinsic motivation, autonomy/control and information. Rewards that were tangible (such as money) appeared to change the perception of what was controlling behavior. The rewards led the participants to think that the reason for which they were engaged in the task was in order to receive the monetary payoff. In conditions without such rewards, the participant would theoretically perceive their own intrinsic motivation as the reason for engaging in the task. The second aspect of this task involved the informational aspect of rewards. Deci suggested that the verbal rewards were likely not behaviorally distinguishable from the participants' own feelings of accomplishment at successfully completing the task.

These results from early research in CET seem to shed some light on the earlier posed question of why the extrinsic rewards may have not had a positive effect on performance in the earlier described studies (Clark et al., 2012; Kang & Pashler, 2012). Several CET studies suggest that monetary rewards undermined a participants' intrinsic motivation, which may be why Kang and Pashler's incentive was not successful in the first two studies. In addition, these studies suggest that intangible rewards may not undermine intrinsic motivation, which might explain why offering students less time succeeded in increasing their final performance in some of the conditions.

Since this early CET research seems to have illuminated two of the three aforementioned mysteries concerning the experiments about motivation and retrieval enhanced learning, it is possible that another related theory could also aid in other unanswered questions, such as the involvement of personality factors.

Self-Determination Theory of Motivation

The same theorists and researchers who developed CET developed a larger macro-theory of motivation called Self-Determination Theory (SDT). Ryan and Deci (2000) propose that the purpose of SDT research has been described as "facilitating optimal functioning of natural propensities" (p. 68). With that end in mind, SDT posits five mini-theories; CET, Organismic Integration Theory (OIT), Basic Psychological Needs Theory (BPNT), Goal Contents Theory (GCT), and Causality Orientations Theory (COT) (Deci & Ryan, 2000). Although all five of the SDT mini-theories are related to the current study, COT has particular relevance because it includes variables which are at the level of personality and have not been accounted for in the earlier mentioned experiments concerning motivation and retrieval-enhanced learning.

Whereas CET focuses primarily on facilitating intrinsic motivation, OIT is primarily concerned with conditions that facilitate extrinsic motivation. SDT theorists realize that not all behavior can be motivated intrinsically (Ryan, 1995). OIT posits that extrinsic motivation can exist on a continuum of integration between the external motivation and the individual's autonomy, referred to as "regulatory styles" (p. 72; Deci & Ryan, 2000). By this view, behaviors that are intrinsically motivated are also intrinsically regulated. However, an externally motivated behavior may be seen as "integrated regulation" style in which the person extensively identifies the externally motivated behavior with the self. On the opposite side of the spectrum exists the external regulation style in which behaviors

are only performed for the purpose of rewards and punishments (Ryan & Connell, 1989). The overall point of OIT is that the extent to which the individual elects to autonomously identify with the extrinsically motivated behavior may affect the extent to which the individual will autonomously regulate that behavior (Ryan & Deci, 2000).

Deci and Ryan's BPNT may be the main reason why SDT research moved beyond research concerning intrinsic and extrinsic motivation to theorize concerning larger constructs. This theory suggests that basic needs for autonomy, competence, and relatedness are essential psychological needs that are inherently important for optimal human functioning and intrinsic motivation (Ryan, 1995). These basic needs have been found to be related to overall well-being and congruence between personality and behavior (Reis, Sheldon, Gable, Roscoe, & Ryan, 2000; Sherman, Nave, & Funder, 2011).

Similar to BPNT, another mini-theory within SDT that involves a larger scope than motivation is GCT. In studying the goals which American college students value, Kasser & Ryan (1993) found that there was a negative relationship between students' aspirations toward financial success and their overall well-being. Kasser & Ryan (1996) found similar negative relationships when goals involved seeking social recognition, appealing appearance. GCT has associated these factors that are associated with obtaining the approval of others as "extrinsic." However, GCT has also delineated a set of factors that have been labeled "intrinsic" which have a positive correlation with health and well-being. These goals are personal

growth, affiliation and intimacy, contributing to one's community, and physical health (Ryan, Huta, & Deci, 2008). Niemec, Ryan, and Deci (2009) surveyed students concerning these different goal contents and found that attainment of intrinsic aspirations was positively associated with psychological health, whereas attainment of extrinsic aspirations showed no such relationship.

The fifth and final mini-theory associated with SDT involves individual's tendencies to orient towards different behavioral regulation contexts (Deci & Ryan, 2000). Earlier research in the behaviorist tradition had shown that performance on tasks that were being reinforced markedly increased when the participants perceived the reinforcement as under their control or due to a skill that they could develop (James & Rotter, 1958). In the real world outside of the laboratory, individuals were found to have differences on the extent to which they perceived causality in their lives as due to their internal agency or external sources (Rotter & Mulry, 1965). In COT, Deci and Ryan (1985) posited three separate causality orientations instead of two. The proposed causality orientations were aligned with other important concepts of SDT, named Autonomy, Controlled, and Impersonal. COT may be of particular relevance for the current study because it is an individual characteristic that might interact with different situational conditions of extrinsic rewards (Deci & Ryan, 1985; Hagger & Chatzisarantis, 2011; Koestner & Zuckerman, 1994).

Causality Orientations and Individual Characteristics

After many years of only experimental studies concerning CET, Deci, Nezlek, and Sheinman (1981) suggested that CET related variables may exist as an interaction between the individual giving the reward, the individual receiving the reward, and the situation. To that end, these researchers measured the individual tendencies that teachers and students had in relation to the constructs of autonomy and control. The authors reported that there was a relationship between the characteristics of the teachers and the characteristics of the children. Although this study was not directly concerned with COT, it may have been the first study in SDT research to begin considering individual differences in orientations towards different types of motivation. In the first study published concerning COT, Deci and Ryan (1985) suggested that they had noticed how different individuals seem to react differently to the experimental manipulations that were often experienced by undergraduates in their studies. These individual causality variables seem to be related to personality variables (Deci & Ryan, 1985; Olesen, 2011; Olesen et al., 2010) and real world outcomes such as work and academic achievement (Lam & Gurland, 2008; Ratelle, Guay, Vallerand, Larose, & Sencal, 2007; Wong, 2000).

Deci and Ryan (1985) developed a scale that measured individual causality orientations. The authors theorized that an individual that has a high autonomy causality orientation (ACO) would seek situations that allow for internal regulation of behavior, instead of external rewards and constraints. Having an ACO was also theorized to lead one to view external rewards such as academic honors as indicators of one's competence, rather than ends in themselves. Positive

correlations were found between the constructs of ACO and ego development, self-esteem, and supporting autonomy in children. Negative correlations were found between ACO and self-derogation, private self-consciousness, guilt, and hostility.

Deci and Ryan (1985) also suggested that individuals who are higher on measures of the controlled causality orientation (CCO) would be more focused on shaping their environment through sets of rewards and reinforcements. These individuals were hypothesized to be more likely to see extrinsic rewards as determining their behavior. The CCO scale was found to have positive correlations with measures of having an external locus of control, being publically and privately self-conscious, exhibiting hostility, type-A personality behaviors, and placing higher value on performing admirably. In addition, students with higher CCO ratings were found to underperform on their exams.

Finally, individuals who rate higher on a measurement of the impersonal causality orientation (ICO) were theorized to see themselves as incompetent and unable to reliably regulate their own behavior. These individuals were hypothesized by Deci and Ryan (1985) to believe that their behavior is caused by impersonal forces that are beyond their control. Scores from the ICO scale were positively correlated with measures of social anxiety, self-derogation, depression, public and private self-consciousness, hostility, fear, shame, guilt, and having an external locus of control.

Since these causality orientations tended to have relationships with many different personality variables, it seems possible that there could be redundancy

between these and other personality variables. Oleson et al. (2010) utilized factor analysis methods to assess the level of redundancy amongst the causality orientations and the big five personality characteristics that have been often used in psychology research (openness, conscientiousness, extraversion, agreeableness, and neuroticism; for review, see Poropat, (2009)). Using the same measure of causality orientations as the previously mentioned Deci and Ryan (1985) study, items known to measure ACO were found to load on a separate factor from the measured personality traits. The items related to CCO loaded onto one particular distinct factor, but not as well as the ACO items. Olesen et al. suggested that CCO could have some overlap with the personality characteristic of agreeableness. ICO also loaded heavily on a distinct factor, but shared some cross-loading items with neuroticism. Although the causality orientations measures seemed to be distinct from personality traits, Olesen et al. found relationships among these variables. The measures of ACO showed strong positive correlations with extraversion and openness. Weaker correlations were found between ACO and agreeableness and conscientiousness. Olesen et al. reported a negative correlation between CCO measures and agreeableness. The measure of ICO was found to have a significant positive correlation with neuroticism and a negative correlation with extraversion. In a subsequent study, a different sample was used as well as confirmatory factor analysis and latent modeling to confirm these factor distinctions and correlations (Olesen, 2011).

As mentioned earlier, several researchers have studied the relationship between causality orientations and academic achievement. Wong (2000) conducted a long-term study on the relationship between causality orientations and achievement amongst a selected group of talented high school students. These students completed causality orientation surveys and carried around a pager for a week. When the pager signaled their attention, they were supposed to answer a set of questions concerning their activities at the moment. Using this methodology, the measure of ACO was positively related to ease of concentration, perceived control and competence, doing work that is of importance to the self, and the experience of flow while doing schoolwork. Wong reported that ACO was not related to academic achievement in this study. The measure of CCO was positively related to doing activities that were of importance to others and level of alertness or activation. CCO had a negative relationship with academic achievement. Wong also investigated the combination of the constructs of ACO and CCO. Students who reported higher ratings on the CCO measure and low-to-moderate ratings on the ACO measure tended to take lower level academic courses than students who had higher ratings on the ACO measure. This final type of analysis suggests that these different types of causality orientations may not be cut and dried distinct concepts, but may be more context sensitive.

Ratelle et al. (2007) used cluster analyses to investigate the motivational profiles of different students with respect to their causality orientations and level of education. When samples of high school students were involved, three main

profiles were found. Ratelle et al. reported that these students either had higher CCO ratings, moderate ratings of both CCO and ACO, or high ratings of both CCO and ACO. High school students with higher levels of both ACO and CCO tended to have positive school outcomes such as academic achievement and lower absenteeism. When a sample of college students was studied, Ratelle et al. found different profiles were detected in the cluster analysis. The college sample included profiles of students with high ACO, high ACO and CCO, and lower ratings on both ACO and CCO. There were no differences reported in academic achievement across the clusters with higher ACO or higher ACO and CCO. However, Ratelle et al. reported that the profile with higher ACO had a higher tendency to persist in completing an academic program. The authors suggested that these differences in causality orientation profiles might have been indicative of the contexts in which different profiles may thrive. Students in high school may have not reported higher levels of ACO because of the high school environment might not have been supportive of autonomy.

Research concerning causality orientations has not been confined to the situation of the classroom. Lam and Gurland (2008) studied self-determination and employment outcomes of community college employees with different causality orientations. Employees who rated themselves higher on measures of ACO were likely to have higher scores on a measure of intrinsic reasons for performing at work. In turn, both of these measures were found to predict job satisfaction and commitment. By contrast, Lam and Gurland reported that employees who self-

reported as higher on a measure of CCO reported less intrinsic motivation for performing at work. Although their study highly supports the main facets of COT, Lam and Gurland suggested that further studies should investigate the interaction between these individual causality orientations and individual situational contexts.

Causality Orientations and Motivation Contexts

In addition to being valid personality constructs, Deci and Ryan (1985) suggested that causality orientations would be associated with environmental events and elicit similar behaviors and cognitions amongst individuals within these categories. Causality orientations have been found to moderate the effects of feedback (Koestner & Zuckerman, 1994) and the effects of rewards on motivation (Hagger & Chatzisarantis, 2011).

Koestner and Zuckerman (1994) investigated the interaction between causality orientations and motivation during performance on a word maze task. Half of the participants were randomly assigned to a condition in which they were given negative feedback concerning their performance whereas the other half was given positive feedback. For the individuals with an ACO, the feedback condition did not have an impact on their subsequent intrinsic motivation. Individuals with a CCO showed a large discrepancy in subsequent motivation, with the group who received negative feedback reporting higher intrinsic motivation than the group who received positive feedback. Koestner and Zuckerman also reported that the individuals with a higher ICO rated their intrinsic motivation higher if they had

received positive feedback. Therefore it seems that causality orientations may moderate the effect of feedback on motivation.

In a recently published study, Hagger and Chatzisarantis (2011) utilized a design similar to many of the previous studies concerning intrinsic motivation and extrinsic rewards to study the moderating effect of causality orientations (Deci, 1971). The undergraduates in the study were randomly assigned to conditions in which they worked on puzzles with or without receiving a monetary reward for completing them successfully. Intrinsic motivation was measured similarly to the previous studies, as time spent on the puzzles when the experimenter left the room. The deviation from the design of the previously reviewed studies involved the recruitment of the participants in this study. Initially, the participants' causality orientations were assessed. Students who received higher scores on the ACO measure were compared to students who received higher scores on the CCO measure. Hagger and Chatzisarantis reported that receiving a reward for puzzle completion did not decrease intrinsic motivation for the ACO participants. However, CCO participants who received the reward spent less of their free time on the puzzles than the CCO participants who did receive the reward. In comparing the groups across conditions, the ACO participants who received the reward spent more time on the puzzles than the CCO participants who received the reward. Similarly, the ACO participants who did not receive the reward spent more time on the puzzles than the CCO participants that did not receive the reward.

As hypothesized by Deci, Nezlek, and Sheinman (1981), it appears from these studies that motivation does occur as an interaction between individual personality characteristics, task properties, and situational contexts. The individual characteristics of causality orientations (autonomy, controlled, impersonal) interact with task properties (intrinsically motivating or extrinsic; (Koestner & Zuckerman, 1994), both of which interact with situational incentives (reward vs. no reward; (Hagger & Chatzisarantis, 2011).

In conclusion, many years of research have been devoted to the understanding of different learning and motivational processes that simultaneously exist within an educational environment (Pintrich, 2003; Roediger, 2008). Retrieval-enhanced learning has been an effective tool in improving recall, transfer, and subsequent learning while reducing forgetting (Carpenter et al., 2008; Pastotter et al., 2011; Roediger & Karpicke, 2006a; Rohrer et al., 2010). Several different contexts and groups of individuals also appear to benefit from retrieval-enhanced learning (Gates, 1922; McDaniel et al., 2011; Tse et al., 2010). Self-determinations theory posits that motivational variables such as intrinsic motivation, psychological needs, and causality orientations also impact behavior in educational environments (Deci, 1971; Deci & Ryan, 1985, 2000).

Preliminary evidence from two studies suggests that it is unclear whether motivational moderate the effect of retrieval on learning (Clark et al., 2012; Kang & Pashler, 2012). Therefore, many questions remain to be answered. The current study has been designed to shed light on the following research questions.

1. How do extrinsic incentives moderate the effect of retrieval on learning?

2. Does intrinsic motivation moderate the effect of retrieval on learning?

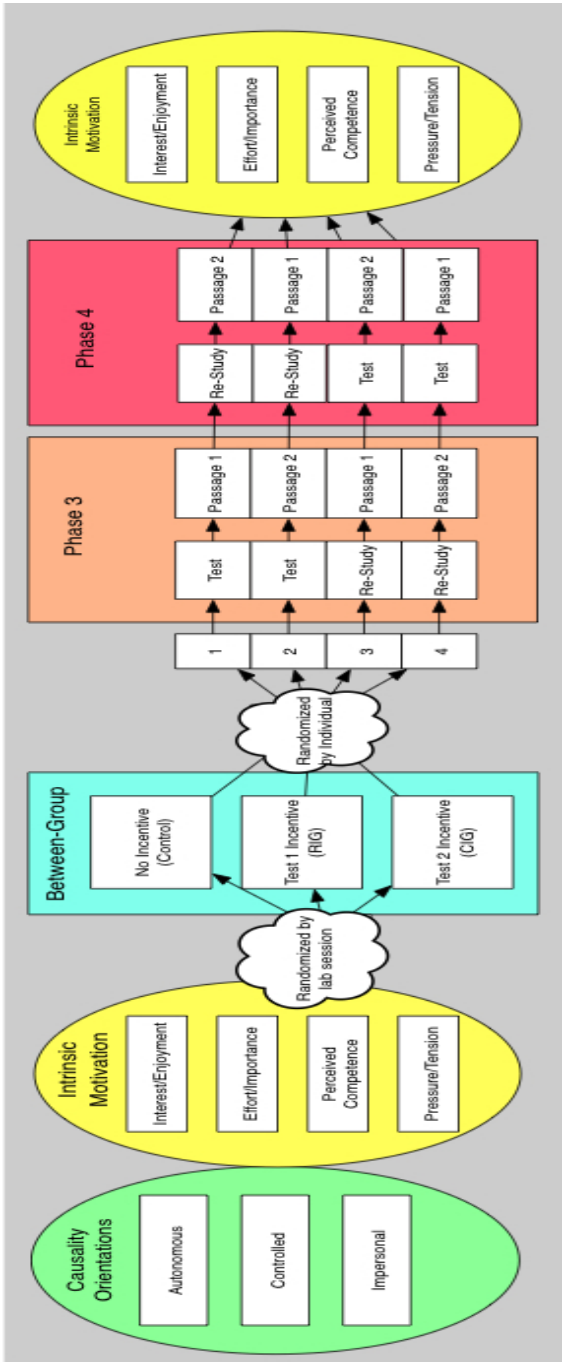
3. Do individual causality orientations moderate the effect of retrieval on learning?

Knowing the answer to these questions could help educators and instructional designers to understand if using retrieval as a learning intervention would particularly helpful given the motivation of the learners involved.

Chapter 3. Method

In the current study, some elements of previous studies, especially Clark et al. (2012) and Kang and Pashler (2012), have been incorporated along with new elements from Self-Determination Theory. Since the constructs of motivation and retrieval-enhanced learning have been studied through different methods, the current study combines these elements both statistically and methodologically (see figure 1 for a graphical overview).

Figure 1. Overall Study Design



Participants

The participants were 154 undergraduate students assigned to participate in research studies by the Educational Psychology subject pool. Students in this pool were drawn from a set of undergraduate educational psychology courses that represent a wide array of subject matter. They tend to mirror the general characteristics of the undergraduate population of the institution.

A power analysis was conducted using G*Power 3.1 in order to determine the number of participants that would be necessary to achieve statistical significance. Using a standard effect size of $f^2 = .15$ and 20 predictors, G*Power suggested that 157 participants would be necessary to achieve a statistical power level of .8.

Instruments

Causality measurement. The General Causality Orientations Scale (GCOS) was developed to assess "relatively enduring aspects of people that characterize the source of initiation and regulation" (Deci & Ryan, 1985: p. 109). Three different orientations were assessed; Autonomy orientation (ACO), controlled orientation (CCO), or impersonal orientations (ICO). The format of this scale involves the presentation of 12 different vignettes, such as:

> "A close (same-sex) friend of yours has been moody lately, and a couple of times has become very angry with you over 'nothing.' You might:"

Following each vignette, three different possible responses to such a situation are presented, each of which reflects a particular causality orientation, such as "Share your observations with him/her and try to find out what is going on for him/her" which reflects an Autonomous Orientation. After reading each statement, the participants rate the likelihood that they would engage that particular response in that situation. A previous study has used this measure and found it to have adequate reliability (around .75) as well as predictable correlations to many other constructs theorized to be similar (Deci & Ryan, 1985).

Intrinsic motivation for the task

Modified portions of the Intrinsic Motivation Inventory (IMI) was used for this study. This inventory was designed to measure "participants' subjective experience related to a target activity" (Ryan, 1982; p. 405). The format of this instrument involves a set of randomized statements for participants to endorse on a 7-point Likert-type scale. Lower numbers on the scale indicate that the participants view the statement as "not at all true," whereas higher numbers suggest that these statements are "very true." Some of the items (7) within this instrument are reverse coded. The authors of this instrument have noted that it is particularly suited for modification to fit into particular situational constraints for research purposes. McAuley, Duncan, and Tammen (1989) found that the IMI had a decent reliability overall ($\alpha = .85$).

The overall IMI may involve as many as 7 scales and 45 items, yet the version used for this study included 5 scales and 30 items. The 5 scales were interest/enjoyment

(α = .78), perceived competence (α = .80), effort/importance (α = .84), pressure/tension (α = .85), and value/usefulness (α = not reported). Example items in the interest/enjoyment subscale (I/ES) include: "This activity was fun to do" and "I enjoyed this activity very much." The Perceived Competence Subscale (PCS) included items such as "I was pretty skilled at this activity" and "I am satisfied with my performance at this task." Items that reflected the Effort/Importance Subscale (E/IS) included "I tried very hard on this activity" and "I didn't put much energy into this (reversed)." Example items for the Pressure/Tension Subscale (P/TS) included: "I felt very tense when doing this activity" and "I felt very pressured while doing these." Finally, items that reflected the Value/Usefulness Subscale (V/US) included: "I think this is an important activity" and "I believe doing this activity could be beneficial to me."

The wording of the IMI was modified for both of the times that this scale was administered in this study. During the first administration of Baseline Intrinsic Motivation (B-IMI), the items were oriented toward the task participants were about to do. For example, it said: "I will try very hard on this activity" or "I will be very nervous about doing this activity." During the second administration of this scale (henceforth referred to as the S-IMI, for Study Intrinsic Motivation Inventory), the wording was oriented toward the individual's assigned study condition, such as "I tried very hard to re-study this material" or "I tried very hard on this test."

Experimental Design

In order to answer the questions posed, a design that incorporated both experimental manipulations and self-report survey instruments was chosen. An extrinsic incentive was offered to some students as the primary between-groups variable, whereas study methods was varied within groups.

Between group manipulation: The between-group independent variables manipulated were 1) an incentive for participation and 2) the timing of the administration of the incentive (See Figure 1). All of the participants individually watched a short instructional video which administered this manipulation. For the participants who received an incentive, the incentive was similar to the one used in a previous study, in which the participants were told that their performance would reduce the amount of time required for them to complete their study participation (Clark et al., 2012).

Within group manipulation: In addition to the between-group experimental manipulation, there was a within-group variable that was manipulated. The within-group variable was the different study methods by which the participants reviewed the material: retrieval in the form of answering short-answer questions (the retrieval condition) or re-studying. Each participant experienced both methods but the order in which they occur was randomized to control for order effects. Also, the order with which participants saw the two types of content was randomized to control for order effects (see figure 1). The computer software that was used for this study will perform participant randomization automatically. This type of randomization has been done in previous studies concerning the effect of retrieval (Butler & Roediger, 2008).

Manipulation fidelity tasks: Several survey instruments were administered at different times in this study. These survey instruments measured the fidelity of the experimental manipulations from the perspective of the participants. For example, did a participant understand the condition under which he was responding?

Materials

A set of two short passages that were used in a previous study (A. Butler & Roediger, 2008) were used as content for this experiment. Written permission to use these materials was obtained from the first author of the previous study. Passage one has 552 words and is about the artist Salvador Dali. Passage 2 has 602 words and is about a war that occurred between the Israelis and several Arab nations in 1973.

When the participants were practicing retrieval, they answered 10 short answer questions concerning the content of the passage. These questions ranged from factual content, such as "What was the name of Salvador Dali's wife?" to more conceptual questions such as "What was the main difference between Salvador Dali's public and private lives?" After answering each of these questions, the participants were given feedback by being shown the correct answer to the question by the computer software. In the follow-up session the next week, all of the participants answered the same 10 questions in a multiple-choice format.

When the participants were re-studying, they were shown the stems and the answers from the retrieval questions in the form of a statement. This same procedure was used in a previous study (Butler & Roediger, 2008). For example, one of the questions

asks. “In what region of Spain was Salvador Dali born?” and the answer is “Catalonia.” For this question, the participants who are re-studying this content read a sentence that said; “Salvador Dali was born in the town of Figueres in the Catalonia region of Spain.” Time constraints were added to the re-study condition so that the participants had to view each page for at least 5 seconds and no more than 30 seconds. These time constraints served to keep the participants from simply clicking through the re-studying content as well as keeping them from spending an inordinate amount of time on the content, which could have created a disparity between the time spent by the testing and re-studying conditions.

Experimental Manipulations

Control Group. The participants randomly assigned to the control group did not receive any instructions to suggest that their performance on the measures in the study would have any consequence. This group watched a video that reminded them that they needed to return for the second part of the study in order to receive credit at the conclusion of the study.

Experimental Groups. For the remaining participants (2/3 of the total) watched a short video which gave them an incentive for performing well.

Retrieval Incentive Group (RIG). The incentive for this group was reduced testing time based on their performance on the upcoming task. Specifically this group (1/3 of the total participants) watched a video which told them that their performance on some of the questions they were going to answer during the first

session would count towards reducing the amount of time they were required to participate during the final session. The software would also inform them prior to reading the passages:

> "The content of this passage is important!
>
> If you remember this content and receive a particular score on the questions you answer in a few minutes, you will not have to stay and complete follow-up questions during next week's session!"

They were also instructed that the software would indicate which particular questions counted towards reducing participation time. Before the RIG participants answered the initial retrieval short-answer questions, a page of instructions in big red letters informed them:

> "This next group of 10 questions is important!
>
> If you receive a particular score on these questions, you will not have to stay and complete follow-up questions during next week's session!"

There was also a reminder on the page of each short-answer question they answered during the retrieval section. During the re-study portion, they were given these instructions on screen:

> "This next group of 10 statements is important!
>
> If you remember this content well, you may not have to stay and complete follow-up questions during next week's session!"

However, when they returned for the second session a week later, they were not reminded of this particular incentive or given any warnings by the software.

During the second session, the RIG participants were simply directed towards the task in the same way as the control group.

Criterion Incentive Group (CIG). The incentive for this group was reduced testing time based on their performance on the final test the following week. Specifically, these participants watched a video that instructed them that their performance on some of the questions they would answer during the final session of the experiment would count towards reducing the amount of time they would be required to participate during the final session. Prior to reading the passages, they were given this instruction:

> "The content of this passage is important!
>
> If you remember this content next week and receive a particular score on the questions you take next week, you will not have to stay and complete follow-up questions during next week's session!"

During the second session, the CIG was informed prior to taking the criterion multiple-choice test by a page of similar instructions in big red letters that performing well on the following questions could reduce the amount of time required for their participation. There was also a reminder on the page of each multiple-choice question they answered during the retrieval section.

Procedures

This study consisted of five separate phases (See figure 2).

Phase 1: In the first phase, the experimenter explained the study and requested the participants complete an informed consent document as per institutional IRB

requirements. After consenting, the participants began by responding to the items of the GCOS and the B-IMI.

Phase 2: The second phase included the manipulation of the incentive. All of the participants watched a short video on their individual computers. There were three different (but similar) versions of this video that were viewed by the different groups. For the participants randomly assigned to the control group, the video was simply a reminder for them to return. The participants in the RIG and the CIG saw videos which explained their incentive for performing well.

Phase 3: The participants began phase three by reading the first passage they were assigned. The first passage read could be either passage one or two dependent upon which condition they were assigned (see figure 1). After reading the first passage, the participants engaged in one of the interpolating tasks for about five minutes to create a short break between reading the material and the study method.

> The first interpolating task that was chosen to create a buffer time period between reading the material and engaging in the study method was a test of working memory. This test is very similar but not identical to the letter-number sequencing measure that is included in the Wechsler Intelligence scales. It involves the participants watching short videos in which a set of letters and numbers appear for one second at a time. After the video, the participants move to the next page and record the letters they remember in alphabetical order. Next, they move to another page and record the numbers they remember in numerical order. Although the reliability of this particular method is not known, this construct was included

because working memory was found to moderate the effect of retrieval in a previous study (Brewer & Unsworth, 2012). This task was not included in the analysis for this study; it was included for two other purposes: 1) to create a time gap between the reading and the study method, and 2) to engage the working memory in another task. Previous studies showed that a retrieval intervention is not as effective when it occurs from working memory or immediately after the initial encoding (Karpicke & Roediger, 2007; Roediger, 2008).

Following the interpolating task, the participants studied the material using retrieval short answer questions (*retrieval)* or reading the questions stated as sentences (*re-study)*. After completing their respective study method, the participants completed the S-IMI to test the fidelity of the incentive manipulation, specifically to measure the intrinsic motivation experienced by the participants during the task.

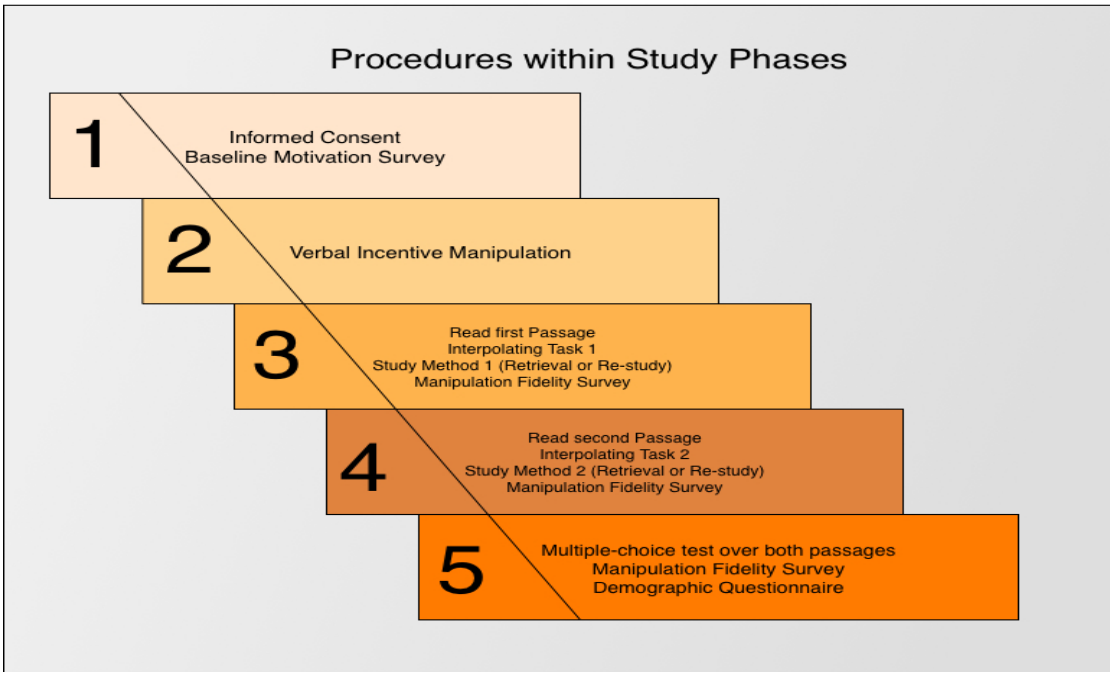
Phase 4: Then the participants moved on to phase four, which involved completing this same process again with a different reading passage, cognitive buffer task, and study method. In the analyses of the data acquired, the different conditions were combined so that each group had individuals who were randomized into the different conditions of the order of the passages and the study methods.

The second interpolating task used to create a buffer between reading and studying was Raven's Progressive Matrices (Raven, 1938). This task involves the participant being shown an image that has a part missing. The participant is also shown a list of possible images that could complete the incomplete image. Finally, the participant is instructed to choose from the set of options which image best completes the incomplete

image. This measure was used in a prior study and it was also found to be related to the effect of retrieval on learning (Brewer & Unsworth, 2012). Similar to the first interpolating task, this measure was also not part of the planned analyses.

Phase 5: The fifth phase occurred one week after the first session. Upon arrival to the session, the participants in the RIG and the CIG were reminded about the incentive that was given to them during the prior week verbally and by the computer software. The control groups were simply directed to take the final assessment in order to complete their participation. The participants also completed another S-MI after they finished the questions. Finally, the participants completed demographic questionnaires and were debriefed regarding the purpose of the study.

Figure 2. Study Procedures



As mentioned previously in the literature review, there were three main research questions to be discussed in this study. Each question was broken down into two or three hypotheses. The next chapter describes the analyses related to each of these questions.

Chapter 4. Results

Question 1: Do extrinsic incentives moderate the effect of retrieval-enhanced learning?

Hypothesis 1A: It was predicted that the group which received an extrinsic incentive for performance on the final criterion task would benefit from the retrieval study method less than the group that did not receive an incentive.

A mixed-model ANOVA was conducted to assess the effects of extrinsic incentives and study methods on learning. The within-subjects variable was study method (re-study vs. retrieval). The between-subjects variable was incentive (none vs. retrieval incentive vs. criterion incentive).

There was a main effect of study method ($F(1, 152) = 74.87, MSE = 1.25, p < .001$), where retrieving resulted in better performance than re-studying. For the main effect of incentive, an a priori contrast was conducted comparing the two incentive groups combined vs. the control group with no incentive. In this comparison, the combined incentive group scored higher than the control group, $t(153) = 2.03, p < .05$. In the analysis that specifically addressed Hypothesis 1A, there was no interaction between incentive and the study method, $F(1, 153) = 1.12, MSE = 1.25, p > .3$. This suggests that retrieval was equally effective across the incentive and control groups.

Hypothesis 1B. The group that received an incentive on the initial retrieval task was predicted to benefit more from retrieval than the group that received the incentive on the final task.

To address this question, data from only from the two groups receiving incentives were analyzed. Again, a mixed-model ANOVA was conducted to assess the separate effects of 1. type of incentive and 2. retrieval and study methods as well as the interaction between the two. Again the within-subjects variable was study method (re-study vs. retrieval) and the between-subjects variable was type of incentive (retrieval incentive vs. criterion incentive). Again there was a main effect of retrieval, $F(1, 103) = 51.43$, $MSE = 108.76$, $p < .001$, where retrieving resulted in better performance than re-studying. There was no main effect of incentive type ($F(1, 103) = 0.06$, $MSE = 3.56$, $p = .80$). There was also no interaction between retrieval and incentive type, ($F(1, 103) = 1.64$, $MSE = 108.76$, $p = .20$). Thus, there was no difference between initial and final incentives on the retrieval effect.

Table 1. Analysis Results for Research Questions 1 and 2

Analysis Results for Research Questions 1 & 2

Hypothesis	Source	df	*F*	*MSE*	*p*
		Within Subjects			
	Study Method	1, 152	74.87	1.25	.001**
	Study Method x Incentive	1, 152	1.12	1.25	.329
		Between Subjects			
	Incentive	1, 152	2.71	2.08	.069*
1B		Within-Subjects			
	Study Method	1, 103	51.43	1.05	.001**
	Study Method x Incentive	1, 103	1.64	1.05	.203
		Between Subjects			
	Incentive	1, 103	0.06	3.55	.802
2A		Within-Subjects			
	Study Method	1, 153	74.39	1.25	.001**
	Study Method x Intrinsic Motivation	1, 153	1.53	1.25	.218
		Between Subjects			
	Intrinsic Motivation	1, 153	7.45	4.09	.007**
2B		Within-Subjects			
	Study Method	1, 153	69.88	1.25	.001**
	Study Method x Intrinsic Motivation	1, 153	1.15	1.25	.286
	Study Method x Incentive	1, 153	0.86	1.25	.354
	Study Method x Incentive x Intrinsic Motivation	1, 153	.067	1.25	.796
		Between Subjects			
	Intrinsic Motivation	1, 153	9.62	3.94	.002**
	Incentive	1, 153	6.07	3.94	.015*
	Incentive x Intrinsic Motivation	1, 153	1.75	3.94	.187

*p <.10, **p < .01

Question 2: Does intrinsic motivation moderate the effect of retrieval on learning?

Hypothesis 2A. It was predicted that participants with lower intrinsic motivation for the task would benefit more from retrieval enhanced learning.

To address this question, a mixed-model ANOVA was conducted to assess the separate effects of intrinsic motivation and retrieval interventions and the possible interaction between these two. The within-subjects variable was study method (re-study vs. retrieval). The between-subjects variable was intrinsic motivation (low vs. high). The participants were split into two groups based on the median value of self-reported scores on the intrinsic motivation inventory. Again there was a main effect of retrieval, $F(1, 153) = 74.39, MSE = 190.47, p < .001$, where retrieving resulted in better performance than re-studying. There was also a main effect of intrinsic motivation, $F(1, 153) = 0.06, MSE = 4.09, p < .01$, where those with high intrinsic motivation performed better than those with low intrinsic motivation. There was no interaction between retrieval and intrinsic motivation, $F(1, 153) = 1.53, MSE = 190.47, p = .22$. Thus, students with low intrinsic motivation did not benefit from retrieval more than those with high intrinsic motivation. The planned hierarchical linear regression was also performed with similar results. The self-reported intrinsic motivation for the task and subject matter was not related to the effect of retrieval on learning ($r^2 = .002, F(1, 153) = .34, MSE = 2.50, p > .5$).

Hypothesis 2B. As predicted by SDT, intrinsic motivation was expected to have a different relationship with retrieval-enhanced learning dependent upon the extrinsic incentive condition. Within the conditions that did not receive an incentive, intrinsic motivation would be negatively related to retrieval-enhanced learning. Within the conditions that received an incentive, intrinsic motivation would be positively related to retrieval-enhanced learning.

Using an analysis similar to hypothesis 2A, a mixed-model ANOVA was conducted to assess the combined effects of motivational factors (extrinsic and intrinsic), retrieval as a learning intervention, and the interactions. The within-subject factor was study method (restudy vs. retrieval). The between subjects factors included level of intrinsic motivation (low vs. high) and incentive group (incentive vs. non-incentive). There was a main effect of retrieval, $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. There was also a main effect of intrinsic motivation, $F(1, 151) = 9.62$, $MSE = 3.94$, $p < .01$, where those with higher intrinsic motivation performed better than those with lower intrinsic motivation. There was also a main effect of the extrinsic incentive, $F(1, 151) = 6.07$, $MSE = 3.94$, $p < .02$, where those who received an extrinsic incentive performed better than those who did not. Although these main effects were observed, there were no two-way significant interactions between study method and extrinsic incentives, $F(1, 151) = .86$, $MSE = 1.25$, $p > .3$, study methods and intrinsic motivation, $F(1, 151) = 1.14$, $MSE = 1.25$, $p > .28$, or different types of motivation (extrinsic incentives and intrinsic motivation), $F(1, 151) =$

1.75, $MSE = 3.94$, $p > .18$. No three-way interaction was observed between study method, extrinsic incentives, and intrinsic motivation ($F(1, 151) = .06$, $MSE = 1.25$, $p > .79$).

Similar to the previous hypothesis, the planned hierarchical regression analysis was conducted. The results indicated that including the dummy-coded variables for the incentive groups did not result in a substantial increase in the ability of the model to predict the retrieval effect score ($r^2 = .016$, $F(2, 152) = .80$, $MSE = 2.50$, $p > .4$). The third block containing the interaction terms for the respective groups also did not improve the prediction above the effect of chance ($r^2 = .018$, $F(2, 149) = .54$, $MSE = 2.53$, $p > .7$). Together, these results suggest that hypothesis 2B was not supported by the data. It does not appear that the effect of intrinsic motivation differed according to the incentive that was administered.

Table 2. Analysis Results for Hypothesis 3A

Analysis Results for Hypothesis 3A

Causality Orientation					
Autonomy	Source	df	*F*	*MSE*	*p*
		Within Subjects			
	Study Method	1, 153	73.43	1.26	.001**
	Study Method x Autonomy	1, 153	0.12	1.26	.730
		Between Subjects			
	Autonomy	1, 153	1.94	4.23	.166
Controlled		Within-Subjects			
	Study Method	1, 153	70.74	1.24	.001**
	Study Method x Controlled	1, 153	1.84	1.24	.177
		Between Subjects			
	Controlled	1, 103	0.34	4.28	.559
Impersonal		Within-Subjects			
	Study Method	1, 153	71.21	1.26	.001**
	Study Method x Impersonal	1, 153	0.28	1.26	.600
		Between Subjects			
	Impersonal	1, 153	2.76	4.22	.099*

*p <.10, **p < .01

Question 3: Do causality orientations moderate the effect of retrieval on learning?

Hypothesis 3A. It was predicted that overall, causality orientations would not significantly predict retrieval-enhanced learning.

Mixed-model ANOVAs were conducted to assess the combined effects of the causality orientations and the different study methods. For these analyses, study method was the within-subject variable. The between-subject variable was the scores on the three self-reported causality orientations scales (autonomy, controlled and impersonal with a low vs. high split similar to the previous analyses on each scale).

The first analysis involved the *autonomy* orientation. This analysis revealed a main effect of retrieval, $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. However, higher ratings of the autonomy causality orientation scale were not associated with better performance ($F(1, 151) = 1.94$, $MSE = 8.22$, $p > .15$). No interactions were observed between these variables ($F(1, 151) = .12$, $MSE = 1.25$, $p > .73$).

The second analysis involved the *controlled* orientation. This analysis revealed a main effect of retrieval, $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. However, higher ratings of the controlled causality orientation were not associated with better performance, $F(1, 151) = .34$, $MSE = 4.28$, $p > .5$. No interaction was observed between these variables, ($F(1, 151) = .12$, $MSE = 1.25$, $p > .73$).

The third analysis involved the *impersonal* orientation. This analysis revealed a main effect of retrieval $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. However, ratings on the impersonal causality orientation had no effect on performance, $F(1, 151) = 2.76$, $MSE = 4.21$, $p = .099$. No interaction was observed between these variables, $F(1, 151) = .27$, $MSE = 1.25$, $p > .59$.

Finally, the planned hierarchical linear regression was conducted in order to consider if the causality orientations were related to the effect of retrieval enhanced learning. The dependent variable for this particular analysis was also the retrieval effect score. The individual's ratings of the separate causality orientations were grand-mean centered and entered as the predictors in the first block of the analysis. As predicted by hypothesis 3A, the first block of the analysis did not substantially improve the prediction of the retrieval effect score ($r^2 = .025$, $F(3, 151) = 1.3$, $MSE = 2.48$, $p > .2$). This result suggests that the individual ratings of causality orientations were not related to the effect of retrieval on learning.

Hypothesis 3B. As in previous research (Hagger & Chatzisarantis, 2011), the three causality orientations would have effects in interactions with the incentive conditions.

A three-way mixed-model ANOVA was conducted to assess the combined effects of the Autonomy causality orientations, extrinsic incentives, and the different study methods. Study method was the within-subjects factor. The between-subject variables

were the scores on the self-reported autonomy causality orientations scale (low vs. high) and the extrinsic incentives.

The first analysis involved the autonomy scale orientation relationships to other variables. This analysis revealed a main effect of retrieval, $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. A main effect on performance was also observed for the participants who received an extrinsic incentive, $F(1, 151) = 4.33$, $MSE = 4.11$, $p < .05$. However, higher ratings of the autonomy causality orientation were not associated with better performance, $F(1, 151) = 1.94$, $MSE = 8.22$, $p > .15$. No two-way interactions were observed between study method and autonomy orientation, $F(1, 151) = .52$, $MSE = 1.25$, $p > .47$, or study method and incentive ($F(1, 151) = 1.09$, $MSE = 1.25$, $p > .29$). There was also no three-way interaction between these variables ($F(1, 151) = .86$, $MSE = 1.25$, $p > .35$).

The second analysis involved the controlled orientation. This analysis revealed a main effect of retrieval, $F(1, 151) = 69.89$, $MSE = 189.23$, $p < .001$, where retrieving resulted in better performance than re-studying. A main effect on performance was also observed for the participants who received an extrinsic incentive ($F(1, 151) = 4.33$, $MSE = 4.11$, $p < .05$). However, higher ratings of the controlled causality orientation were not associated with better performance ($F(1, 151) = .34$, $MSE = 4.28$, $p > .5$). No two-way interactions were observed between study method and controlled orientation, $F(1, 151) = 1.23$, $MSE = 1.25$, $p > .26$, or study method and incentive, $F(1, 151) = 1.13$, $MSE = 1.25$, $p > .28$. There was also no three-way interaction between these variables ($F(1, 151) = .45$, $MSE = 1.25$, $p > .5$).

The third analysis involved the impersonal orientation. This analysis revealed a main effect of retrieval, $F(1, 151) = 67.51$, $MSE = 1.26$, $p < .001$, where retrieving resulted in better performance than re-studying. A main effect on performance was also observed for the participants who received an extrinsic incentive ($F(1, 151) = 4.73$, $MSE = 4.13$, $p < .05$). Higher ratings on the impersonal causality orientation had no main effect on performance, ($F(1, 151) = 2.38$, $MSE = 4.13$, $p > .12$). No two-way interactions were observed between study method and impersonal orientation, $F(1, 151) = .41$, $MSE = 1.26$, $p > .51$, or study method and incentive ($F(1, 151) = .95$, $MSE = 1.26$, $p > .33$). There was also no three-way interaction between these variables ($F(1, 151) = .02$, $MSE = 1.26$, $p > .8$).

Finally, the planned hierarchical linear regression was also conducted. . As in the previous analyses, including variables associated with the causality orientations did not improve the prediction ($r^2 = .04$, $F(3, 151) = 1.3$, $MSE = 2.47$, $p > .2$). The interaction terms between the incentive groups and the centered causality orientation variables also did not improve the prediction ($r^2 = .08$, $F(3, 151) = 1.23$, $MSE = 2.45$, $p > .2$). This result suggests that the interaction between the incentive condition and the self-reported causality orientation measures was not related to the effect of retrieval on learning.

In general across all the analyses, retrieval as a study method continued to make a positive impact on the learners across all the comparisons. The presence of incentives or intrinsic motivation had a positive effect on performance. However, there were no interactions between the motivational variables and the effect of retrieval.

Table 3. Analysis Results for Hypothesis 3B

Analysis Results for Hypothesis 3B

Causality Orientation					
Autonomy	Source	df	*F*	*MSE*	*p*
			Within Subjects		
	Study Method	1, 151	73.30	1.26	.001**
	Study Method x Autonomy	1, 151	0.52	1.26	.471
	Study Method x Incentive	1, 151	1.09	1.26	.298
	Study Method x Incentive x Autonomy	1, 151	0.87	1.26	.354
			Between Subjects		
	Autonomy	1, 151	2.69	4.11	.103
	Incentive	1, 151	4.34	4.11	.039*
	Autonomy x Incentive	1, 151	1.68	4.11	.197
Controlled			Within-Subjects		
	Study Method	1, 151	67.01	1.25	.001**
	Study Method x Controlled	1, 151	1.24	1.25	.268
	Study Method x Incentive	1, 151	1.13	1.25	.289
	Study Method x Incentive x Controlled	1, 151	0.45	1.25	.503
			Between Subjects		
	Controlled	1, 151	0.52	4.18	.559
	Incentive	1, 151	5.60	4.18	.019*
	Controlled x Incentive	1, 151	0.0	4.18	.998
Impersonal			Within-Subjects		
	Study Method	1, 151	67.51	1.26	.001**
	Study Method x Impersonal	1, 151	0.28	1.26	.600
	Study Method x Incentive	1, 151	0.95	1.26	.519
	Study Method x Incentive x Impersonal	1, 151	0.03	1.26	.876
			Between Subjects		
	Impersonal	1, 151	2.38	4.13	.125
	Incentive	1, 151	4.74	4.13	.031*
	Impersonal x Incentive	1, 151	0.36	4.13	.548

*p <.10, **p < .01

Chapter 5. Discussion

There is overwhelming empirical evidence that using retrieval is an effective means of learning in several different contexts. There is also evidence for several different motivational variables that affect learning. However, there have been no studies that have investigated how motivation may moderate the effectiveness of retrieval as a learning intervention. Clark et al. (2012) suggested that increasing motivation might reduce the effectiveness of retrieval as a learning intervention, whereas Kang et al. (2012) suggested that increasing motivation might improve learning overall without moderating the effect of retrieval. The current study examined the role of extrinsic incentives, intrinsic motivation, and causality orientations with regard to the retrieval effect. The results of this study suggest that retrieval is an effective learning intervention for students regardless of these motivational factors.

The Effect of Incentives

Previously, Clark et al. (2012) found that participants who received an extrinsic incentive to perform well on a retrieval task surprisingly performed worse on a final long-term test than those who did not receive an incentive. Therefore, it was expected that the participants in the current study who received an extrinsic incentive might benefit less from using retrieval when compared to participants who were not given an extrinsic incentive. Not surprisingly, as in many previous studies, using retrieval as a study method led to better performance than simply re-reading statements concerning the content. Also

rather unsurprisingly, participants who received an extrinsic incentive to perform well outperformed the participants who did not receive an extrinsic incentive. Contrary to the predicted outcome (hypothesis 1A), the extrinsic incentive did not moderate the effect of retrieval on learning. Retrieval was equally effective across the different incentive conditions. One possibility for this failure to find moderation could be related to the differing types of questions and feedback utilized. Short answer items were used in this study because previous studies have found that short-answer retrieval tasks are generally better for learning (Fazio et al., 2010). However, the previous study (Clark et al., 2012) utilized multiple-choice retrieval items.

Another difference was the feedback mechanism utilized. In the previous study, the computer would provide corrective feedback in which both the correct answer and the participants' answer would be displayed at the same time. This feedback would inform the participant if the chosen answer was indeed correct. The correct answer would also be marked, regardless of whether or not it was chosen by the participant. In the current study, feedback was given by displaying the correct answer immediately after the participant answered the question.

Previous studies have also suggested that increased effort during retrieval may improve learning (Bjork, 1994; Pyc & Rawson, 2009). Therefore, hypothesis 1B predicted that receiving an incentive for performing well on the initial retrieval questions would invoke more effort and result in better performance, as opposed to receiving an incentive for performing well on the final criterion questions. The current results did not

support the prediction of hypothesis 1B. There were no differences in performance between the groups who received incentives at different times in the process.

Intrinsic Motivation and the Effect of Retrieval

SDT suggests that there is much more to motivation than extrinsic incentives for performance; sometimes individuals are intrinsically motivated to perform certain tasks (Deci & Ryan, 2000). Because motivation through extrinsic incentives has been shown to affect retrieval interventions in previous studies (Clark et al., 2012, Kang et al., 2012), it was predicted that intrinsic motivation might have a similar effect. In previous studies, individuals with lower performance on ability tests (such as working memory and IQ) benefitted from retrieval more than those with higher performance on ability tests (Brewer & Unsworth, 2012). Therefore, it was predicted that individuals with lower intrinsic motivation would benefit from retrieval more than individuals with higher intrinsic motivation (hypothesis 2A). In the current study, the participants who self-reported higher levels of intrinsic motivation did indeed perform better on the outcome learning measures than the participants with lower self-reported intrinsic motivation. However, higher intrinsic motivation appears to have benefitted both learning conditions equally. There was no interaction between study method (retrieval vs. restudy) and level of intrinsic motivation (high vs. low). Hypothesis 2A was not supported by the data. One possible explanation for not supporting this prediction was that intrinsic motivation might only create an interaction between these variables within certain conditions, such as when the participants were not receiving an extrinsic incentive.

Many studies using the framework of SDT found that offering an extrinsic incentive for a task had the effect of reducing intrinsic motivation (Deci, 1971, 1972; Deci & Ryan, 1999). Remembering the previous study which involved lower scoring participants gaining more benefits from retrieval (Brewer & Unsworth, 2012), it was expected that intrinsically motivated participants who did not receive an extrinsic incentive would benefit less from retrieval than participants with less intrinsic motivation (Hypothesis 2B). However, hypothesis 2B also predicted that participants who received an extrinsic incentive would show the opposite pattern. As mentioned previously, both receiving an extrinsic incentive and having higher intrinsic motivation led to better performance in this study. The effect of intrinsic motivation did not change according to the extrinsic incentive. Individuals with higher levels of intrinsic motivation performed better regardless of whether an incentive was offered. The effect of retrieval on learning was not moderated by extrinsic incentives, intrinsic motivation, or the interaction between the two. Hypothesis 2B was not supported by the data in this study. Another possibility that remains is that the effect of these motivation-related variables could hinge on personality variables.

Causality Orientations and the Effect of Retrieval

Causality Orientation Theory suggests that individuals may prefer different types of motivational environments (Deci & Ryan, 1985). Some individuals prefer environments where they can operate autonomously, free of controlling rewards or punishments. Others may prefer a controlled environment, where rewards and

punishments are clearly delineated. Finally, some individuals may have an impersonal orientation where they feel that there is no relationship between their actions and rewards or punishments. In the current study, these causality orientations were not expected to moderate the effect of retrieval on learning (hypothesis 3A). This prediction was based on a previous study in which causality orientations moderated the effect of rewards, not actual motivation (Haggar & Chatzisarantis, 2011). Consistent with the prediction, participants' rating on the autonomy causality orientation scale did not moderate the effect of retrieval on learning. Higher ratings on the controlled causality orientation also did not moderate the effect of retrieval on learning. Finally, the effect of retrieval on learning was not moderated by ratings on the impersonal causality orientation scale. Therefore, hypothesis 3A was supported by the data. Although these causality orientation variables did not moderate the effect of retrieval, the previous study (Haggar & Chatzisarantis, 2011) suggested that they moderate the effect of rewards on performance.

Haggar & Chatzisarantis (2011) found that individuals who rated higher on an autonomy causality orientation scale did not suffer a reduction in motivation from receiving a reward, while participants with a higher controlled causality orientation did experience a reduction in motivation. The evidence from that study suggests that the effect of an extrinsic incentive may be contingent upon the causality orientation of the individual receiving the incentive. It was predicted in this study that the interaction between individual causality orientation and extrinsic incentives would moderate the effect of retrieval on learning (hypothesis 3B). Although participants who received an

extrinsic incentive performed better than those who did not, ratings on the causality orientation scales did not moderate the impact of the incentive or the effect of the retrieval intervention. Overall, these results do not support hypothesis 3B or the results of the previous study concerning the moderating effect of causality orientations (Haggar & Chatzisarantis, 2011). It is possible that differences in recruitment could be responsible for the discrepancy. Haggar and Chatzisarantis (2011) specifically chose individuals who rated themselves as higher on one of the causality orientation scales relative to the other. That recruitment method was not utilized in the current study. It seems likely that these causality orientation variables may only moderate the effect of incentives in some situations that bear a resemblance to the situation that was examined in the previous study. These results may have implication that may serve to refine the different theories associated with causality orientations.

Theoretical Implications

Several theories have been discussed in this dissertation concerning outcomes of the current study. These theories have been discussed in the context of the effects of retrieval and motivation on learning.

Spreading Activation Theory, Transfer Appropriate Processing Theory, and Disuse Theory have attempted to account for many of the effects of retrieval. Spreading Activation Theory suggests that attempting to recall information activates numerous traces of memory, increasing the likelihood that the correct information will be recalled (Anderson, 1983). Transfer Appropriate Processing Theory suggests that retrieval

improves memory via an increase in the coherence between the study conditions and the conditions of the final assessment (Morris et al., 1977). Disuse Theory suggests that memory traces involve separate storage and recall traces. These recall traces are thought to be susceptible to being crowded out if they are not successfully recalled (Bjork & Bjork, 1992). In addition, Self-Determination Theory has been discussed with regard to motivation. SDT suggests (among many, many other tenets) that intrinsic motivation improves performance, extrinsic incentives may undermine intrinsic motivation, and causality orientations may interact with situational characteristics such as rewards and motivation. Although the current study was not expressly designed to assess the value of these theories, the results of the study may have implications for these different theoretical perspectives.

In the current study, retrieval had a positive effect on long-term memory. Therefore, the results did not expressly discredit any of the theories concerning the benefit of retrieval. However, one effect was observed that did not necessarily adhere to the perspective of Spreading Activation Theory. As mentioned earlier, several studies have suggested that increased effort in retrieving information may improve the effect of retrieval (Kang et al., 2007; Kornell, Hays, & Bjork, 2009; Pyc & Rawson, 2009). Therefore, one could presume that an intervention which increased effort would also increase the effect of retrieval. In the current study, an incentive improved overall long-term learning, yet it did not improve the effect of the retrieval intervention. Therefore, it could be possible that the incentive may have had the effect of spreading activation, but

in a way that was somehow parallel to the effect of retrieval. Therefore, the results of the current study appear to be more coherent with Disuse theory.

Disuse theory suggests that memory traces may have more than one pathway. One trace involves the storage of the memory, while another involves the memory's ability to be recalled. Bjork & Bjork (1992) suggested that viewing an item more than once may increase the storage trace of that memory, while retrieval increases both the storage and retrieval traces. Using this perspective, it is possible that motivation (intrinsic or extrinsic) from the current study had the effect of strengthening the storage trace of the memory, which is why the effect of motivation occurred similarly across both the re-study and retrieval study methods. According to Disuse theory, using retrieval during study improves the separate retrieval trace of a memory, which also accounts for the positive effect of retrieval which occurred from using retrieval in this study. These separate traces may have enabled the parallel effects of retrieval and motivation that were found in the current study.

Similar to a few other studies, the current study did not fully support the perspective of transfer appropriate processing. Even though the conditions of retrieval were not extraordinarily similar to the conditions upon the final learning assessment, retrieval continued to improve learning. During the retrieval phase, participants answered questions by entering text into a blank text box. The final assessment phase involved choosing the correct answer from a list of possibile answers. Similar results have been found in previous studies, which have suggested that it may be the approximation of the process of recall which accounts for the improvement (Kang, 2007).

The effects of motivation in the current study also did not clearly mirror previous research from an SDT perspective. SDT has many different mini-theories which attempt to explain motivation. Cognitive Evaluation Theory suggests that intrinsic motivation leads to better performance (Deci, Koestner, & Ryan, 1999). In concordance with CET, the current study concluded that students with more intrinsic motivation remembered more of the content. However, CET also suggests that extrinsic incentives may undermine intrinsic motivation. The current results did not support this aspect of CET. Similar to other studies outside of the SDT literature (Cameron, 2001), incentives had the effect of improving performance in the current study. One possible reason for this result could be that the materials were not specifically chosen because of their entertainment value. Several studies and meta-analyses have shown that incentives improve performance when tasks are less interesting in comparison to performance on interesting tasks (Cameron & Pierce, 1994; Murayama & Kuhbandner, 2011; Pierce, Cameron, Banko, & So, 2003).

Limitations

In a study in which the majority of the hypotheses were not supported, it is not surprising that some limitations may have been encountered. Statistical conclusion limitations were encountered due to the measurements that were used in the study. For example the questions from the passage concerning Salvador Dali proved to be less difficult than the questions concerning the 1973 Arab-Israeli conflict (See appendix C for passages and questions).

External validity limitations may also have existed concerning the effect of the incentive that was used in this study. Previous research has suggested that retrieval is not always superior to re-reading material for everyone (Bouwmeester et al., 2011). In the current study, however, it seems that the effectiveness of retrieval may have been moderated by the content that was used. When participants used retrieval to study the passage about Salvador Dali and re-read the statements concerning the 1973 Arab-Israeli war (conditions 1 and 4, see figure 1), the discrepancy between their retrieval and re-reading scores was much higher than when the situation was reversed. Since these were the only dependent measures used in this study, it is difficult to assess the effects of this problem.

It seems probable that the conclusions reached in this study based on the different group interventions should not have been affected by this limitation. Since the participants in the differing conditions were randomly assigned to the different incentive groups, the effects were likely evenly distributed across the groups. In other words, each of the different incentive groups had an equal chance of having participants with the higher retrieval effect scores (conditions 1 and 4 from figure 1) and participants with lower retrieval effect scores (conditions 2 and 3). Although the participants were randomly distributed, it is also possible that the within-group variance in the respective groups was inflated by these differing conditions. Since the retrieval effects were lower in conditions 2 and 3, including all of the conditions in the same group may have increased the within-group variance. Individuals who may have had a larger retrieval effect may have been affected by that condition, reducing the possibility of finding these individual

and group differences. Therefore, the likelihood of making a type II error in these analyses may have been inflated.

Since the differing conditions had different effects of retrieval, it is possible that individuals who had very similar characteristics (such as level of intrinsic motivation) may have incurred very different scores on the respective study methods. This problem may have inflated the type II error for the analyses of hypotheses 2A, 2B, 3A, and 3B. All of these concerns may have had a substantial effect on reliability of the statistical conclusions made based on this study, but there were also concerns regarding the external reliability of the procedure itself as well.

In order to offer an extrinsic incentive for some of the groups in this study, some of the students were told that better performance would reduce the amount of time that they would spend in completing their required study participation. This incentive method was chosen primarily because the current study did not have funding to pay the participants or offer anything else of monetary value. A secondary reason was that it has been used to successfully impact performance in previous studies (Clark et al., 2012; Kang et al., 2012). Although this incentive method was effective in improving the performance of the participants in the current study, it is possible that other types of incentives (such as verbal praise or physical elements) may have had different effects, or no effects at all (Henderlong, 2002). Another possibility could be that the incentive offered in this study could have created a focus on preventing the loss of time, whereas another type of incentive that focused on obtaining something of value (or with a

promotion focus) may have also had different effects than the effects that were observed in this study (Grimm et al., 2008).

Implications for Future Research

One of the unique aspects of the current research study was integrating two different areas of research with the goal of understanding some area of practice. Many studies have shown that retrieval affects learning, and motivation affects learning. Unfortunately, the interaction between these different types of variables has not enjoyed such a long tradition of study. The outcome of the current study may help us to understand why these interaction-type studies have not been more prominent.

The data from the current study were consistent with many of the past studies, which found that retrieval had a positive effect on learning. Similarly, the data also supported that incentives and intrinsic motivation improved learning. These results were similar to a previous study, which was also investigating the effect of incentives on retrieval as a learning intervention (Kang et al., 2012). However, these results failed to replicate another previous study (Clark et al., 2012) in which there was an interaction between incentives and retrieval as a learning intervention. Therefore, future studies are needed to understand the difference between the current results and those of the previous studies. For example, it could be that the method of feedback had a negative effect on the participants who were concerned about their performance (due to the incentive), whereas the method of feedback used in the current study did not have such an effect. Future

research could also investigate whether the difference was related to the level of success that the participants had while retrieving the information.

Several previous studies have suggested that increased levels of effort during retrieval may benefit the effectiveness of using retrieval methods (Kornell, Hays, & Bjork, 2009, Pyc & Rawson, 2009). However, the current study did not support the conclusions of the previous studies. In the previous studies, level of effort was manipulated by altering the content to make it difficult (Pyc & Rawson) or impossible (Kornell, Hays, & Bjork) to learn. The current study investigated effort by offering an incentive and including the effects of self-reported intrinsic motivation. Both of these variables had positive effects on performance, but neither of them moderated the effect of retrieval on learning. Therefore, future research should investigate this apparent discontinuity. It seems possible that the current incentive and motivation variables were not strong enough to produce a difference that could moderate the effectiveness of retrieval, whereas the previous studies used procedures designed to more forcefully enhance the effort required from participants.

One important implication for future research would be using multiple types of content for each study method. For example, it would have been better in the current study to have at least two separate passages per study method. Including an additional passage per study method would reduce the likelihood that some of the discrepancy between the different study methods could have been due to differences in the participants' feelings toward the content.

Many of the previous studies reviewed in this dissertation have found that using retrieval as a study method is effective in improving learning across many different contexts. The results from the current study concur with these previous studies in suggesting that the effect of retrieval appears to be a robust technique in which the effect may not change according to the student's level of external or internal motivation.

Appendix A: The General Causality Scale

Directions:

These items pertain to a series of hypothetical sketches. Each sketch describes an incident and lists three ways of responding to it. Please read each sketch, imagine yourself in that situation, and then consider each of the possible responses. Think of each response option in terms of how likely it is that you would respond that way. (We all respond in a variety of ways to situations, and probably most or all responses are at least slightly likely for you.) If it is very unlikely that you would respond the way described in a given response, you should circle answer 1 or 2. If it is moderately likely, you would select a number in the mid range, and if it is very likely that you would respond as described, you would circle answer 6 or 7.

1. You have been offered a new position in a company where you have worked for some time. The first question that is likely to come to mind is:

    a. What if I can't live up to the new responsibility?

    1 2 3 4 5 6 7

    b. Will I make more at this position?

    1 2 3 4 5 6 7

    c. I wonder if the new work will be interesting.

    1 2 3 4 5 6 7

2. You have a school-age daughter. On parents' night the teacher tells you that your daughter is doing poorly and doesn't seem involved in the work. You are likely to:
    a. Talk it over with your daughter to understand further what the problem is
    b. Scold her and hope she does better
    c. Make sure she does the assignments, because she should be working harder

3. You had a job interview several weeks ago. In the mail you received a form letter which states that the position has been filled. It is likely that you might think:
    a. It's not what you know, but who you know
    b. I'm probably not good enough for the job
    c. Somehow they didn't see my qualifications as matching their needs.

4. You are a plant supervisor and have been charged with the task of allotting coffee breaks to three workers who cannot all break at once. You would likely handle this by:
    a. Telling the three workers the situation and having them work with you on the schedule.
    b. Simply assigning times that each can break to avoid problems.
    c. Find out from someone in authority what to do or do what was done in the past.

5. A close (same-sex) friend of yours has been moody lately, and a couple of times has become very angry with you over "nothing." You might:
    a. Share your observations with him/her and try to find out what is going on for him/her
    b. Ignore it because there's not much you can do about it anyway.
    c. Tell him/her that you're willing to spend time together if and only if he/she makes more effort to control him/herself

6. You have just received the results of a test you took, and you discovered that you did very poorly. Your initial reaction is likely to be:
    a. "I can't do anything right," and feel sad.
    b. "I wonder how it is I did so poorly," and feel disappointed
    c. "That stupid test doesn't show anything," and feel angry

7. You have been invited to a large party where you know very few people. As you look forward to the evening, you would likely expect that:
    a. You'll try to fit in with whatever is happening in order to have a good time and not look bad.
    b. You'll find some people with whom you can relate.
    c. You'll probably feel somewhat isolated and unnoticed.

8. You are asked to plan a picnic for yourself and your fellow employees. Your style for approaching this project could most likely be characterized as:
    a. Take charge, that is, you would make most of the major decisions yourself.
    b. Follow precedent: you're not really up to the task so you'd do it the way it's always been done before.
    c. Seek participation: get inputs from others who want to make them before you make the final plans.

9. Recently a position opened up at your place of work that could have meant a promotion for you. However, a person you work with was offered the job rather than you. In evaluating the situation, you're likely to think:
    a. You didn't really expect the job; you frequently get passed over.
    b. The other person probably "did the right things" politically to get the job.
    c. You would probably take a look at factors in your own performance that led you to be passed over.

10. You are embarking on a new career. The most important consideration is likely to be:
    a. Whether you can do the work without getting in over your head.
    b. How interested you are in that kind of work.
    c. Whether there are good possibilities for advancement.

11. A woman who works for you has generally done an adequate job. However, for the past two weeks her work has not been up to par and she appears to be less actively interested in her work. Your reaction is likely to be:
    a. Tell her that her work is below what is expected and that she should start working harder.
    b. Ask her about the problem and let her know you are available to help work it out.
    c. It's hard to know what to do to get her straightened out.

12. Your company has promoted you to a position in a city far from your present location. As you think about the move you would probably:
    a. Feel interested in the new challenge and a little nervous about it at the same time
    b. Feel excited about the higher status and salary that is involved.
    c. Feel stressed and anxious about the upcoming changes.

Appendix B: Intrinsic Motivation Inventory

Directions:
For each of the following statements, please indicate how true it is for you, using the following scale:

Not at all true Somewhat True Very True

1 2 3 4 5 6 7

1. I think I will enjoy reading the passage very much
2. Reading and learning can be fun to do.
3. I think this will be a boring activity.
4. This activity will probably not hold my attention at all.
5. This activity might be interesting.
6. I think this activity will be quite enjoyable.
7. Before reading these passages, I am thinking I could enjoy it.
8. I think I am pretty good at reading and answering questions.
9. I think I can do pretty well at this activity compared to other students.
10. After working at this activity for a while, I felt pretty competent.
11. I think I can perform to my satisfaction at reading and answering questions.
12. I am pretty skilled at reading and answering questions.
13. I do not do very well at reading and answering comprehension questions. (R)
14. I am planning on putting a lot of effort into this.
15. I will not have to try very hard to do well at this activity. (R)
16. I will try very hard on this activity.
17. It is important to me to do well at this task.
18. I am not going to put much energy into this. (R)
19. I am not at all nervous about doing this. (R)
20. I will probably feel very tense while doing this activity
21. I will be relaxed while doing these. (R)
22. I might get anxious while reading this passage and answering the questions.
23. I will feel pressured while doing these.
24. I believe this activity could be of some value to me.
25. I would be willing to do this again because it has some value to me.
26. I believe doing this activity could be beneficial to me.
27. I think this is an important activity.
28. I think it is important for college students to participate in research such as

this one.

29. I am glad to participate in research studies as a service to the University.
30. I resent that I have to participate in research for credit in my courses. (R)
31. I would like for the researcher in this study to benefit from my participation.
32. I am planning on doing as little as possible in order to get credit for this study. (R)
33. I would participate in more research studies such as this for free if someone asked me.
34. I would like to complete my participation in this research as quickly as possible, even if it results in my participation not being of value to the study.

Appendix C: Passages and Questions

Passage 1

Salvador Felipe Jacinto Dalí, known popularly as Salvador Dalí, was a Spanish artist. One of the most important painters of the 20th century, he created some of the most widely recognized images that came out of the Surrealist movement in the visual arts. His best-known work, *The Persistence of Memory*, was completed in 1931, while he was living in the Montparnasse quarter of Paris.

In this famous piece, Dalí introduced the image of the soft, melting pocket watch in an attempt to debunk the assumption that time is rigid or deterministic. Dalí wanted to express pure psychic automatism, the real functioning of thought in the absence of all control exercised by reason. His paintings are easily identified by their striking, bizarre, and dreamlike images.

Before he became a renowned artist, Dalí grew up in the town of Figueres in the Catalonia region of Spain with parents who were strict disciplinarians, but also very unconventional. As a result, his childhood was rather unusual and marked by unique events that would shape his personality. For example, Dalí's parents treated him as though he was the reincarnation of his older brother, also named Salvador, who had died of meningitis prior to the artist's birth.

As he grew into adulthood, Dalí developed an affinity for doing unusual things to draw attention to himself. He grew a giant, pencil-thin moustache, which became iconic of him, and often wore flamboyant outfits, including a purple cape and cane. An artist of great imagination, Dalí's behavior sometimes irked those who loved his art as much as it annoyed his critics, since his eccentric manner sometimes drew more public attention than his artwork.

During the middle period of career, the politics of Salvador Dalí played a significant role in his emergence as a successful artist. In his youth, Dalí embraced for a time both anarchism and communism, but eventually became a supporter of Francisco Franco's fascist regime. Some of Dalí's public statements supported the repression enacted under Franco's reign and praised Franco for signing death warrants for political prisoners.

Still, it is impossible to determine whether his tributes to Franco were sincere or whimsical. Indeed, some of Dalí's private actions signal that his views were somewhat different. He continued praise of Federico García Lorca even in the years when Lorca's works were banned and he was a friendly acquaintance of famed architect and designer Paul Laszlo, who was ethnically Jewish.

Late in his career, Dalí did not confine himself to painting but experimented with many unusual or novel media and processes. In 1960, Dalí began work on the Dalí Theatre and Museum in his hometown of Figueres, a project that would not be completed until 1974. He also spent much of his time with his wife and muse, Gala, whom he had married in 1934, before her sudden death in 1982.

The final years of Dalí's life were marked by two possible suicide attempts, first in 1983 when he deliberately dehydrated himself and then in 1984 when a fire broke out in his bedroom under unclear circumstances. After these events, Dalí was rescued by a group of his friends who saw to it that he was comfortable living in his Theater-Museum for his final years. He died of heart failure at Figueres on January 23, 1989 at the age of 84.

Passage 1 Short-Answer Questions

1. Salvador Dali created some of the most widely recognized images to come out of what artistic movement?

2. In what region of Spain was Salvador Dalí born?

3. For what action did Salvador Dalí praise Francisco Franco?

4. What was the name of Salvador Dalí's wife?

5. What image did Salvador Dali portray in The Persistence of Memory?

6. What was a major influence on Salvador Dalí's eccentric personality?

7. What was the main difference between Salvador Dali's public and private political views?

8. Why might Salvador Dalí have attempted to commit suicide?

9. Dali praised the work of Federico Garcia Lorca even though his works were…

10. What was the name of Dali's brother who died of Meningitis?

Passage 1 Multiple-Choice Questions

1. Salvador Dali created some of the most widely recognized images to come out of what artistic movement?
   - The Surrealist movement
   - The Futurist movement
   - The Fauvist movement
   - The Art Nouveau movement
   - The Symbolist movement

2. In what region of Spain was Salvador Dalí born?
   - He was born in the region of Catalonia
   - He was born in the region of Andalucia
   - He was born in the region of Aragon
   - He was born in the region of Cantabria

- He was born in the region of Valencia

3. For what action did Salvador Dalí praise Francisco Franco?
   - Signing death warrants for political prisoners
   - Imposing martial law after capturing Madrid
   - Proclaiming Spain a monarchy, but not designating a monarch.
   - Banning the usage of any language other than Spanish
   - Expelling members of the rival Carlist political party

4. What was the name of Salvador Dalí's wife?
   - Her name was Gala
   - Her name was Aracelia
   - Her name was Elisa
   - Her name was Inez
   - Her name was Marisa

5. What image did Salvador Dali portray in The Persistence of Memory?
   - The image of the soft, melting pocket watch
   - The image of a dark cluster of ants
   - The image of rocks morphing into each other
   - The image of a broken phone over a plate of eggs
   - The image of women disappearing into the clouds

6. What was a major influence on Salvador Dalí's eccentric personality?
   - His unusual childhood
   - His peers at the Art Academy of Madrid
   - His relationship with fellow artist Federico García Lorca
   - His obsession with Freudian theories
   - The poverty he endured as an art student

7. What was the main difference between Salvador Dali's public and private political views?
   - Publically he supported fascism, but privately he was more open-minded.
   - Publically he denounced Franco, but privately he encouraged him.
   - Publically he posed as a monarchist, but privately he liked communism.
   - Publically he appeared uninvolved in politics, but he was active privately.
   - Publically he embraced Catholicism, but privately he abhorred religion.

8. Why might Salvador Dalí have attempted to commit suicide?
   - He was devastated by his wife's death
   - He was upset about the public's disapproval of his work
   - His only son died in an automobile accident
   - He was very poor and destitute
   - He fought chronic depression most of his life

9. Dali praised the work of Federico Garcia Lorca even though his works were…

- Banned
- Rigid and deterministic
- Facist
- Jewish

10. What was the name of Dali's brother who died of Meningitis?
    - Salvador
    - Figueres
    - Felipe
    - Francisco

Passage 2

The 1973 Arab-Israeli War (also known as the Yom Kippur War or the Ramadan War) was fought from October 6 to October 26, 1973, between Israel and a coalition of Arab nations led by Egypt and Syria. The war was part of the larger Arab-Israeli conflict, an ongoing struggle that has resulted in many battles over the years since 1948. Six years earlier, during the Six-Day War, the Israelis had captured the Sinai region from Egypt.

The earlier war had resulted the Israeli occupation of the Golan Heights, which had been taken from Syria. A shaky peace followed in the years following the Six-Day War and Israel erected massive lines of fortification, strengthening its position in the captured lands. Egypt and Syria both desperately desired the return of the land they had lost and this desire was a primary cause of the 1973 Arab-Israeli War.

As a result of this desire to regain lost territory, Egypt and Syria initiated the conflict, starting the war with a surprise joint attack on the day of Yom Kippur. Both the armies cross their respective cease-fire lines and breached Israeli defenses by using water cannons to blast away large barricades built primarily from sand. The Israeli battalion garrisoning the front line was vastly undermanned because of Yom Kippur, a holiday during which most Jews fast and abstain from any use of fire, electricity, engines, communications, etc.

Interestingly, the war also coincided with the Muslim holiday of Ramadan, meaning that many of the Arab soldiers were also fasting. The Israeli forces were quickly overwhelmed because the entire country of Israel had come to a complete standstill and much of the army had been demobilized. The Egyptian and Syria forces had succeeded in catching them by surprise.

The success of the initial attack led to disastrous first week of the war for the Israelis as subsequent offensives by the Arab forces were also successful at pushing the Israeli army back further. Soon both sides had settled into defensive postures, each hoping for the other side to attack. After several days of waiting, the Egyptian generals had grown impatient with their defensive strategy and wanted to help the Syrians, who had been slowly losing ground to the Israelis.

To ease the pressure on the Syrians, the Egyptian generals ordering an all-out attack, which proved to be a total failure and the Egyptian forces suffered massive causalities. The following

day, October 15, the Israelis launched Operation Abiray-Lev (which translates to "Stouthearted Men"), resulting in a successful counterattack against the Egyptians. These events proved to be a turning point in the war and Israel went on to push both the Egyptian and Syrian forces back into their own territories before end of the war.

On October 22, the conflict came to an end as the Security Council of the United Nations unanimously passed Resolution 338 calling for a cease-fire. The 1973 Arab-Israeli war had far-reaching implications for many nations. The Arab world, which had been humiliated by the lopsided defeat of the Arab alliance during the Six-Day War, felt psychologically vindicated by its string of victories early in the conflict.

This vindication paved the way for the peace process that followed, as well as liberalizations in Egypt's foreign policy. The Camp David Accords, which came soon after, led to normalized relations between Egypt and Israel—the first time any Arab country had recognized the Israeli state. On the Syrian front, shuttle diplomacy by Henry Kissinger eventually produced a disengagement agreement in 1974, based on exchange of prisoners-of-war, Israeli withdrawal to the Purple Line, and the establishment of a UN buffer zone.

Passage 2 Questions

1. Before the 1973 Arab-Israeli War, the broader Arab-Israeli conflict had been going on since what year?

2. What did the Egyptian and Syrian forces use to breach Israeli defenses during the start of the 1973 Arab-Israeli War?

3. Launched on October 15th during the 1973 Arab-Israeli War, Operation "Abiray-Lev" translates to what in English?

4. What was the name of the diplomat who brokered a disengagement agreement between Israel and Syria at the end of the 1973 Arab-Israeli War?

5. The 1973 Arab-Israeli War was incited in part because of what two disputed regions?

6. As a result of Yom Kippur and Ramadan, what were both sides of the conflict doing during the 1973 Arab-Israeli War?

7. What change in strategy by the Egyptians led to a turning point in the 1973 Arab-Israeli War?

8. Why were the early events of the 1973 Arab-Israeli War critical to the later normalization of relations between Egypt and Israel?

9. What international organization passed a resolution that called for a cease-fire between the two nations?

10. What was unique about the Camp David Accords?

Passage 2 Multiple-Choice Questions

1. Before the 1973 Arab-Israeli War, the broader Arab-Israeli conflict had been going on since what year?
   - 1948
   - 1928
   - 1901
   - 1963
   - 1957

2. What did the Egyptian and Syrian forces use to breach Israeli defenses during the start of the 1973 Arab-Israeli War?
   - Water Cannons
   - Mechanized infantry
   - Tanks
   - Cruise missiles
   - Bombs

3. Launched on October 15th during the 1973 Arab-Israeli War, Operation "Abiray-Lev" translates to what in English?
   - Operation "Stouthearted Men"
   - Operation "Protect the Heart"
   - Operation "Brave Lion"
   - Operation "Lightening Strike"
   - Operation "Sword of God"

4. What was the name of the diplomat who brokered a disengagement agreement between Israel and Syria at the end of the 1973 Arab-Israeli War?
   - Henry Kissinger
   - John Foster Dulles
   - Kurt Waldheim
   - Margaret Thatcher
   - James Schlesinger

5. The 1973 Arab-Israeli War was incited in part because of what two disputed regions?
   - The Golan Heights and the Sinai region
   - The West Bank and the Gaza Strip
   - Haifa and Elat
   - Beth Shean and Hadera
   - Nahariyya and Safad

6. As a result of Yom Kippur and Ramadan, what were both sides of the conflict doing during the 1973 Arab-Israeli War?
    - Fasting during the day
    - Praying three times a day
    - Refraining from eating fish
    - Resting during the day
    - Dressing in traditional clothing

7. What change in strategy by the Egyptians led to a turning point in the 1973 Arab-Israeli War?
    - Instead of continuing to sit back defensively, the Egyptian army went on an all-out attack
    - Instead of continuing to occupy Israeli land, the Egyptian army pulled out abruptly
    - Instead of continuing to fight on foot, the Egyptian forces used artillery and planes
    - Instead of attacking Israel from the South, the Egyptian army attacked by water from the
    - Instead of sticking by the Syrian forces, the Egyptian army split off to open a new front

8. Why were the early events of the 1973 Arab-Israeli War critical to the later normalization of relations between Egypt and Israel?
    - The early victories helped Egyptian moral and pride, allowing them to recognize Israel
    - The early victories showed that Israel had a superior military and diplomacy was the only option.
    - The co-occurrence of Muslim and Jewish holidays helped both sides realize they had much in common.
    - The early victories left Israel's defenses in ruins, making diplomacy their only option
    - The long military standoff between the two sides gave time for them to talk

9. What international organization passed a resolution that called for a cease-fire between the two nations?
    - The United Nations
    - NATO
    - The Arab Alliance
    - Abiray-Lev

10. What does the passage say was historically unique about the Camp David Accords?
    - An Arab country recognizing the state of Israel
    - The location of the negotiation
    - Vindication for Egypt
    - Syria exchanged prisoners

Appendix D: Text Material Evaluation Questionnaires

Testing Questionnaire

Directions:
For each of the following statements, please indicate how true it is for you, using the following scale:

Not at all true Somewhat True Very True

1 2 3 4 5 6 7

1. While answering these questions, I was thinking about how much I enjoyed it.
2. I did not feel at all nervous while responding to these questions.
3. These questions did not hold my attention at all.
4. I think I responded to these questions pretty well.
5. I would describe these questions as very interesting.
6. I think I understood these questions very well, compared to other students.
7. I enjoyed answering these questions very much.
8. I felt very tense while answering these questions.
9. These questions were fun to answer.
10. I did not feel nervous at all while responding these questions (R)
11. I felt very tense while answering these questions.
12. I was very relaxed in answering these questions. (R)
13. I was anxious while working on these questions.
14. I felt pressured while responding to these questions.

Re-Studying

Directions:
For each of the following statements, please indicate how true it is for you, using the following scale:

Not at all true Somewhat True Very True

1 2 3 4 5 6 7

3. While I was reading this material, I was thinking about how much I enjoyed it.
4. I did not feel at all nervous while reading these statements.
5. This reading material did not hold my attention at all.
6. I think I understood this material pretty well.
7. I would describe this material as very interesting.
8. I think I understood this material very well, compared to other students.
9. I enjoyed reading this material very much.
10. I felt very tense while reading this material.
11. This material was fun to read.
12. I did not feel nervous at all while reading this. (R)
13. I felt very tense while doing this reading.
14. I was very relaxed in doing these. (R)
15. I was anxious while reading these statements.
16. I felt pressured while reading these statements.

Appendix E: Demographic Questions

Demographic Questions

What is your gender?
- Male (1)
- Female (2)

What is your age in years?

What is your highest level of education?
- Freshman (1)
- Sophomore (2)
- Junior (3)
- Senior (4)
- Graduate Student (5)

What is your current estimated GPA?
- Below 2.5 (1)
- 2.5 - 3.0 (2)
- 3.0 - 3.3 (3)
- 3.3 - 3.5 (4)
- 3.5 - 3.7 (5)
- 3.7 - 4.0 (6)

To the best of your knowledge, what was your SAT score?

Vision Please check one of the below options regarding your vision:
- I have never been diagnosed as having a visual impairment (1)
- I have been diagnosed as having a visual impairment, but I wore glasses or contacts during this study (2)
- I have been diagnosed as having a visual impairment and I did not wear glasses or contacts during this study (3)

Please check one of the below options regarding your reading ability:
- I have never been diagnosed as having a learning disability in the area of reading. (1)
- I have been diagnosed as having a learning disability in the area of reading. (2)
- I have never been diagnosed as having a reading disability, but I have struggled with reading throughout my academic career. (3)

How fast of a reader to you consider yourself?
- Very Slow (15)
- Slow (16)
- Neutral (17)

- Fast (18)
- Very Fast (19)

How skilled do you think you are at comprehending what you read?
- Not Very Skilled (1)
- Below Average (2)
- Average (3)
- Above Average (4)
- Very Skilled (5)

Is English your first language?
- Yes (1)
- No (2)

Do you consider yourself multi-lingual?

How many languages do you speak?

Are you currently learning a foreign language in a class or on your own?

How interested are you in learning a foreign language at some point in your life?

What foreign language training have you had?

If you identify primarily with one ethnicity, which ethnicity do you identify with?
- African American (1)
- Asian American (2)
- Hispanic (3)
- Native American (4)
- White (5)
- More than one of these options or other (you will be asked to specify on the next page) (6)

Since you indicated on the previous question that you identify with more than one ethnicity, Please select all that apply.
- African American (1)
- Asian American (2)
- Hispanic (3)
- Native American (4)
- White (5)
- Other (please specify) (6) _____

References

Abott, E. E. (1909). On the analysis of the factor of recall in the learning process. *Psychological Monographs: General and Applied, 11*(1), 159-177.

Adams, R. H., & Strickland, J. (2012). The Effects of Computer-Assisted Feedback Strategies in Technology Education: A Comparison of Learning Outcomes. *Journal of educational technology systems, 40*(2), 211-223.

Agarwal, P., Karpicke, J., Kang, S., Roediger, H., & McDermott, K. (2008). Examining the testing effect with open-and closed-book tests. *Applied cognitive psychology, 22*(7), 861-876.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior, 22*(3), 261-295.

Balcetis, E., & Dunning, D. (2008). A mile in moccasins: How situational experience diminishes dispositionism in social inference. *Personality and Social Psychology Bulletin, 34*(1), 102.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research*, 89-99.

Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology, 102*(4), 773.

Beebe, R., Vonderwell, S., & Boboc, M. (2010). Emerging Patterns in Transferring Assessment Practices from F2f to Online Environments. *Electronic Journal of e-Learning, 8*(1), 1-12.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. *From learning processes to cognitive processes: Essays in honor of William K. Estes, 2*, 35-67.

Bouwmeester, S., & Verkoeijen, P. P. J. L. (2011). Why do some children benefit more from testing than others? Gist trace processing to explain the testing effect. [doi: 10.1016/j.jml.2011.02.005]. *Journal of Memory and Language, 65*(1), 32-41.

Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language, 66*(3), 407-415.

Butler, A., & Roediger, H. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Learning, Memory, 36*(5), 1118-1133.

Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(4-5), 514-527.

Carpenter, S., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36*(2), 438-448.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563.

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology, 19*(5), 619-636.

Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review, 14*(3), 474-478.

Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review, 13*(5), 826-830.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20*(6), 633-642.

Clark, D., Crandall, J., & Robinson, D. (2012). *The Interaction between Incentives and the Testing Effect*. Paper presented at the The Convention for the Association of Psychological Science.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*(1), 105-115.

Deci, E. L. (1972). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology, 22*(1), 113-120.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627.

Deci, E. L., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology, 40*(1), 1.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. *Advances in experimental social psychology, 13*, 39-80.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*(2), 109-134.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268.

Dickhauser, O., Reinhard, M., Diener, C., & Bertrams, A. (2009). How need for cognition affects the processing of achievement-related information. *Learning and Individual Differences, 19*(2), 283-287.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology*: Teachers college, Columbia university.

Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology, 73*(6), 1224-1237.

Fazio, L., Agarwal, P., Marsh, E., & Roediger, H. (2010). Memorial consequences of multiple-choice testing on immediate and delayed tests. *Memory & Cognition, 38*(4), 407.

Gates, A. I. (1922). *Recitation as a factor in memorizing*: The Science press.

Glover, J. A. (1989). The 'testing' phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology, 81*(3), 392-399.

Greeno, J., Collins, A., & Resnick, L. (1996). Cognition and learning. *Handbook of educational psychology*, 15-46.

Grimes, P. (2002). The overconfident principles of economics student: An examination of a metacognitive skill. *Journal of Economic Education*, 15-30.

Hacker, D., Bol, L., Horgan, D., & Rakow, E. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(1), 160-170.

Hagger, M. S., & Chatzisarantis, N. L. D. (2011). Causality orientations moderate the undermining effect of rewards on intrinsic motivation. *Journal of Experimental Social Psychology, 47*(2), 485-489.

Hammond, W. A. (Ed.). (1902). *Aristotle's Psychology: A Treatise on The Principle of Life*. New York: The Mcmillan Co.

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology, 43*(4), 289-294.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81.

Hsieh, P., Acee, T., Chung, W. H., Hsieh, Y. P., Kim, H., Thomas, G. D.et al. (2005). Is educational intervention research on the decline? *Journal of Educational Psychology, 97*(4), 523.

Izawa, C. (1967). Function of test trials in paired-associate learning. *Journal of Experimental Psychology, 75*(2), 194.

Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology; Journal of Experimental Psychology, 83*(2, Pt. 1), 340.

James, W. H., & Rotter, J. B. (1958). Partial and 100% reinforcement under chance and skill conditions. *Journal of Experimental Psychology, 55*(5), 397-403.

Jensen, P., & Moore, R. (2008). Students' behaviors, grades & perceptions in an introductory biology course. *The American Biology Teacher, 70*(8), 483-487.

Kang, S., McDermott, K., & Roediger, H. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*(4), 528-558.

Kang, S., & Pashler, H. (2012). *Does Motivation Modulate the Benefit of Retrieval Practice?* Paper presented at the Convention for the Association of Psychological Science.

Karpicke, J., Butler, A., & Roediger, H. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own?

Karpicke, J. D., & Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 704.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966-968.

Kasser, T., & Ryan, R. M. (1993). A dark side of the american dream: correlates of financial success as a central life aspiration. *Journal of Personality and Social Psychology, 65*(2), 410.

Koestner, R., & Zuckerman, M. (1994). Causality orientations, failure, and achievement. *Journal of Personality, 62*(3), 321-346.

Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition, 34*(5), 959-972.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. [doi: 10.1016/j.jml.2011.04.002]. *Journal of Memory and Language, 65*(2), 85-97.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 989-998.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*(4), 279-308.

Lam, C. F., & Gurland, S. T. (2008). Self-determined work motivation predicts job outcomes, but what predicts self-determined work motivation? *Journal of Research in Personality, 42*(4), 1109-1115.

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences, 21*(1), 124-128.

Mayer, R. E. (1996). Learners as information processors: Legacies and limitations of educational psychology's second metaphor. *Educational Psychologist, 31*(3-4), 151-161.

McAuley, E. (1989). Psychometric properties of the intrinsic motivation Inventory in a competitive sport setting: a confirmatory factor analysis. *Research quarterly for exercise and sport, 60*(1), 48-58.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition, 39*(3), 462-476.

McDaniel, M., Roediger, H., & McDermott, K. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*(2), 200.

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology, 103*(2), 399.

McDaniel, M. A., & Masson, M. E. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(2), 371.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*(5), 519-533.

Niemiec, C. P., Ryan, R. M., & Deci, E. L. (2009). The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life. *Journal of Research in Personality, 43*(3), 291-306.

Nowell, C., & Alston, R. (2007). I Thought I got an A! overconfidence across the economics curriculum. *The Journal of Economic Education, 38*(2), 131-142.

Olesen, M. H. (2011). General causality orientations are distinct from but related to dispositional traits. *Personality and Individual Differences, 51*(4), 460-465.

Olesen, M. H., Thomsen, D. K., Schnieber, A., & Tnnesvang, J. (2010). Distinguishing general causality orientations from personality traits. *Personality and Individual Differences, 48*(5), 538-543.

Palva, S., & Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends in neurosciences, 30*(4), 150-158.

Pansky, A. (2012). Inoculation against forgetting: advantages of immediate versus delayed initial testing due to superior verbatim accessibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M.et al. (2007). Organizing instruction and study to improve student learning. IES Practice Guide. NCER 2007-2004. *National Center for Education Research*, 63.

Pastotter, B., Schicker, S., Niedernhuber, J., & Buml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 287-297.

Pellegrino, J. W., & Quellmalz, E. S. (2011). Perspectives on the integration of technology and assessment. *Journal of Research on Technology in Education, 43*(2), 119-134.

Phelps, R. P. (2012). The effect of testing on student achievement, 1910-2010. *International Journal of Testing, 12*(1), 21-43.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? [doi: DOI: 10.1016/j.jml.2009.01.004]. *Journal of Memory and Language, 60*(4), 437-447.

Pyc, M. A., & Rawson, K. A. (2011). Why is test restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Sencal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology, 99*(4), 734.

Raven, J. C. (1938). Raven progressive matrices.

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin, 26*(4), 419-435.

Riding, R. (2005). Individual Differences and Educational Performance. *Educational Psychology, 25*(6), 659-672.

Roediger, H. (2008). Relativity of remembering: Why the laws of memory vanished. *Annual Review of Psychology, 59*, 225-254.

Roediger, H., & Karpicke, J. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181.

Roediger, H., & Karpicke, J. (2006b). Test-enhanced learning. *Psychological Science, 17*(3), 249.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27.

Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied; Journal of Experimental Psychology: Applied, 17*(4), 382.

Roediger, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 1155.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39*(5), 406.

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 233.

Rotter, J. B., & Mulry, R. C. (1965). Internal versus external control of reinforcement and decision time. *Journal of Personality and Social Psychology, 2*(4), 598-604.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review; Psychological Review, 103*(4), 734.

Runquist, W. (1983). Some effects of remembering on forgetting. *Memory & Cognition, 11*(6), 641-650.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68-78.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450-461.

Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality, 63*(3), 397-427.

Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality & Social Psychology Bulletin, 22*, 280-287.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*(5), 749.

Ryan, R. M., Huta, V., & Deci, E. L. (2008). Living well: A self-determination theory perspective on eudaimonia. *Journal of Happiness Studies, 9*(1), 139-170.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*(2), 113-142.

Schallert, D. L., & Martin, D. B. (2003). A psychological analysis of what teachers and students do in the language arts classroom. *Handbook of research on teaching the English language arts, 2*, 31-45.

Sederberg, P. B., Gauthier, L. V., Terushkin, V., Miller, J. F., Barnathan, J. A., & Kahana, M. J. (2006). Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage, 32*(3), 1422-1431.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2011). Properties of persons and situations related to overall and distinctive personality-behavior congruence. *Journal of Research in Personality*.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research, 95*(6), 323-332.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher, 18*(9), 8-14.

Snow, R. E. (1994). Abilities in academic tasks. *Mind in context: Interactionist perspectives on human intelligence*, 3-37.

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology, 30*(9), 641.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory & Cognition, 35*(5), 1007-1013.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1392.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements, 2*(4), i-109.

Tse, C.-S., Balota, D. A., & Roediger, H. L., III. (2010). The benefits and costs of repeated testing on the learning of face-name pairs in healthy older adults. *Psychology and Aging, 25*(4), 833-845.

Tse, C.-S., & Pu, X. (2012). The Effectiveness of Test-Enhanced Learning Depends on Trait Test Anxiety and Working-Memory Capacity. *Journal of Experimental Psychology: Applied*.

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of verbal learning and verbal behavior, 6*(2), 175-184.

Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology, 6*(1), 94-102.

Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology, 24*(8), 1183-1195.

Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science, 3*(4), 240-245.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological review, 66*(5), 297.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review, 18*(6), 1140-1147.

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25*(5), 731-739.

Wong, M. M. H. (2000). The relations among causality orientations, academic experience, academic performance, and academic commitment. *Personality and Social Psychology Bulletin, 26*(3), 315-326.